Journal of

CHRONIC DISEASES

VOL. 1 JUNE, 1955 NO. 6

OL.
1
955

Original Communications

RHEUMATIC HEART DISEASE AS A PROBLEM OF PREVENTIVE CARDIOLOGY*

GUNNAR BIÖRCK, M.D.

MALMÖ, SWEDEN

From the Department of Medicine, Allmänna Sjukhuset, Malmö, and the University of Lund, Sweden

(Received for publication Dec. 27, 1954.)

THE incidence of rheumatic fever and rheumatic heart disease is known to show considerable variations in different countries, and also within different parts of the same country. Thus, such diseases are common in the New England States, in Great Britain, and in Southern Italy (Naples), whereas they are rare in Switzerland. In some places it is reported that typical rheumatic fever is uncommon, while rheumatic valvular disease is not.[1] Sweden probably occupies an intermediate position; in its interior these diseases appear to be less frequent than along its western coast.[2]

True statistics of morbidity are, generally, not easily available. When, in 1950, I moved from Stockholm to Malmö, a town of 200,000 inhabitants, located at the coast in the southern part of Sweden, I had a definite impression that rheumatic heart disease and rheumatic fever were more frequent in Malmö than in Stockholm. In the course of later studies, this impression has been re-

*This article is based upon papers read before the Second World Congress of Cardiology, Washington, D. C., and the Finnish Medical Society, Helsingfors, in 1954. The statistical analysis of the material was aided by a grant from the Swedish Medical Research Council.

inforced. In our country, the interest in so-called rheumatic diseases has been largely focused on rheumatoid arthritis, and with few exceptions[3] no systematic efforts seem to have been made to provide victims of rheumatic fever with an adequate follow-up organization. Thus, the approach to the problems discussed in this paper may appear backward to Americans, but simply represents an attempt at catching up with progress already made elsewhere.

When, in 1950, operative treatment for mitral stenosis became available to us, we considered it important to get into contact with such patients in Malmö as might suffer from mitral stenosis in an operable phase. Because we were looking for operable cases, such a survey could not be based on patients with advanced valvular heart disease alone, but had to include also the precursors of such heart disease, viz., rheumatic diathesis.

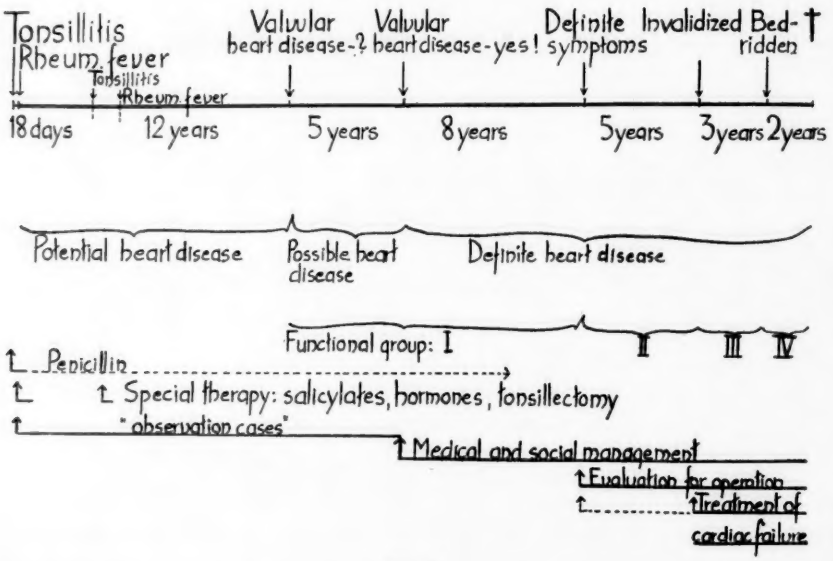


Fig. 1.

In so doing, we were faced with a problem much larger than the one that gave impetus to our efforts. We realized that as cardiologists we were often merely dealing with acute episodes or final products of a chronic disease, with different manifestations and different problems along its clinical course, and we found ourselves challenged to revise our attitude toward the problem in terms of preventive cardiology. This reorientation was strongly supported by the almost simultaneous appearance on the medical scene of three so remarkable happenings as mitral surgery, the anti-inflammatory hormones, and the rheumatic fever prophylaxis program. Finally, there was something else for these patients to do other than move to Arizona or succumb to inevitable cardiac invalidism. To us the situation depicted itself somewhat along the lines that are represented in Fig. 1. (Figures given for time intervals are only suggestions and may vary considerably).[4]

In this scheme, there seemed to be some strategic points, on which efforts ought to be concentrated. The *first* one, obviously, was the adequate treatment of the first attack of rheumatic fever (although with application of the penicillin treatment to every instance of tonsillitis, most attacks of rheumatic fever might be prevented). The *second* point would be the prevention of rheumatic recurrences. *Third* and *fourth* would be the medical control of patients with possible and potential heart disease in order to make a correct diagnosis at an early time. *Last*, a decision as to surgical intervention would have to be made at the proper moment. Naturally, questions of general hygiene, vocational guidance, etc., and of adequate medical treatment would have to be considered all the time; they will, however, not be dealt with here.

Now, what have we learned with regard to these problems from our follow-up study of rheumatic fever patients in Malmö and from our experience with surgery of mitral valvular disease?

OL.
1
955

The follow-up study comprised 1,540 patients out of about 4,000 who had been hospitalized at Malmö General Hospital for rheumatic fever or some equivalents and/or valvular heart disease, diagnosed as rheumatic or congenital, during the years 1930-1950 (Table I). The study was performed in the department of medicine during the year 1953.[5]

TABLE I. SELECTION OF THE MATERIAL

Hospital records from the departments of Medicine, Infectious Diseases, Pediatrics, and Orthopedics during the period 1930-1950 were extracted with regard to patients diagnosed as suffering from:

Rheumatic fever (definite or doubtful)
Chorea minor
Purpura rheumatica
Other possibly "rheumatic" manifestations (such as *erythema exudativum multiforme* and *erythema nodosum*)
Organic heart disease of presumably rheumatic origin
Acute nephritis
Congenital heart disease

The number of patients covered by such records amounts to 4,000. Of these patients, invitations were sent to 1,856 patients whose present address was found. Of these patients, 1,540 (83 per cent) appeared for examination in 1953. Of the rest, some have died in the hospital (approximately 500). These records are now under study. Others either may be alive or have died outside the hospital.

The examination comprised in all cases: analysis of earlier hospital records; history—medical, cardiologic, social; complete physical examination; hemoglobin, blood counts, erythrocyte sedimentation rate, urinalysis; electrocardiogram (4 leads); full-plate x-ray of chest with determination of heart volume; furthermore, if indicated, twelve-lead electrocardiogram, phonocardiogram, or ballistocardiogram; serologic tests—Wassermann, antistreptolysin titer, agglutination of sensitized sheep cells; basal metabolic rate; further hospital study in some cases.

All patients got a letter indicating the result of the examination and giving advice regarding further medical observation. A copy of this letter and a summary of all pertinent findings were put at the disposal of the patient's personal physician.

The material collected during this study is at present subject to statistical treatment. To the data obtained on living patients will be added corresponding data on about 500 patients

who have died in the meantime. This will bring the follow-up to about 50 per cent of the total material for these twenty-one years, which will probably be sufficient to make it representative for the hospitalized population of Malmö, as there is only one hospital in the town.

In answer to the questions posed, Table II confirms earlier investigations[6] and speaks strongly in favor of a program of rheumatic fever prophylaxis.

TABLE II. RECURRENCE OF RHEUMATIC ACTIVITY RELATED TO DEVELOPMENT OF VALVULAR HEART DISEASE (V.H.D.)

NUMBER OF ATTACKS OF DEFINITE RHEUMATIC FEVER (INCLUDING CHOREA)	NUMBER OF CASES	V.H. I	% V.H.D.
One	361	77	21
Two	109	37	34
Three or more	48	22	46

One of the most difficult and most important points in this connection is, of course, the criteria for a diagnosis of rheumatic fever and the extent to which conditions etiologically related to rheumatic fever should be included in a common group. In the first respect, we have applied, with small modifications, the criteria proposed by Jones[7] in 1944. Patients in whom the previous diagnosis of rheumatic fever met these criteria were labeled "rheumatic fever, definite," while the rest of them were labeled "rheumatic fever, uncertain." Cases with chorea *only* were labeled "chorea," while those with concomitant signs of rheumatic fever or with other attacks of rheumatic fever were placed in the rheumatic fever group. Erythema nodosum of probably nontuberculous origin and acute glomerulonephritis—one-half of this group consisting of scarlet fever nephritis—were investigated as examples of other manifestations of hyperergy, presumably to streptococcic agents.

It emerges from Table III that the incidence of valvular heart disease in the two latter groups is very low, and that patients with these diseases, therefore, probably do not need a prophylactic program with regard to the heart.

In the general discussion, as, for example, at the recent Second World Congress for Cardiology, the patient with rheumatic fever is often referred to as "the rheumatic child." This conception, which may have been true a generation ago, does not cover the present situation, as shown by Table IV, indicating the age at the first attack of rheumatic fever, etc., in our material.

It is difficult to say whether this age distribution, particularly in the rheumatic fever group, represents a changing pattern of the age of occurrence of streptococcal infection and development of immunity, coinciding with changes in environmental hygiene, family size, etc., in comparison with earlier years, as seems to be the case with poliomyelitis. It is commonly stated that rheumatic fever nowadays not only is less frequent than before, but also has a milder clinical course. A changing clinical pattern and a greater awareness of somewhat atypical manifestations in adults may play a role in this connection, and

TABLE III. INCIDENCE OF VALVULAR HEART DISEASE (V.H.D.) IN DIFFERENT GROUPS

DIAGNOSIS	SEX	TOTAL	V.H.D.		
			TOTAL	%	%
Rheumatic fever (definite)*	M	196	38	19	25
	F	271	81	30	
Rheumatic fever (uncertain)	M	88	5	6	11
	F	139	20	14	
Chorea minor*	M	13	1	8	33
	F	38	16	42	
Erythema nodosum	M	16	0	0	1
	F	104	1	1	
Acute nephritis	M	154	1	0.6	0.3
	F	157	0	0	
Other manifestations	M	44	3	6	6
	F	45	3	6	

*Patients with chorea *and* other rheumatic manifestations are grouped in the first group. Patients with *only* chorea are grouped in the third group.

TABLE IV. AGE AT FIRST ATTACK OF ACUTE RHEUMATIC FEVER AND OTHER POSSIBLE "RHEUMATIC" MANIFESTATIONS

YEARS		0-4	5-9	10-14	15-19	20-24	25-29	30-34	35-39	40-44	45-49	>50
Rheumatic fever (definite)	M	1	32	45	24	28	29	14	9	5	3	5
	F	6	31	49	43	35	32	27	18	12	7	9
Rheumatic fever (uncertain)	M	4	14	10	10	11	11	6	5	6	2	8
	F	5	8	13	17	22	22	17	9	10	4	9
Chorea	M	—	7	4	1	1	—	—	—	—	—	—
	F	—	15	14	4	1	1	—	—	—	—	—
Erythema nodosum	M	5	7	2	1	1	—	—	—	—	—	—
	F	3	9	11	8	19	17	16	10	5	4	2
Acute nephritis	M	26	61	30	10	4	3	6	4	3	2	5
	F	34	44	30	9	8	13	3	2	6	5	3

make the problems, theoretically and practically, more intriguing and difficult to solve.

As to the question of the best treatment of acute rheumatic fever, we have no statistical material to offer. We believe, at present, salicylates to be the drug for most mild to moderate cases, while cortisone is tried in seriously ill patients with cardiac manifestations of an exudative type.

As for the third and fourth points of the scheme—medical observation and early diagnosis of valvular lesions—the following viewpoints are offered.

In a follow-up study of this kind, some patients undoubtedly will be diagnosed as having valvular heart disease. From the gross diagnostic standpoint they are no problem. The problems arise with the other cases. First, how long a time must elapse before a statement of freedom from valvular heart disease can be given with reasonable safety? And, in the meantime, how often should a medical examination be performed in order to permit an early diagnosis? We have taken the attitude to discharge patients after a period of eighteen to twenty years from their last known manifestation of rheumatic fever if they show no signs or symptoms of valvular heart disease. We have, likewise, recommended control examination every year during the first eight to ten years, from then every second year, and, after fifteen years, once or twice more with three-year intervals. Obviously, with such a policy and a follow-up, dating about twenty-five years back, a very large number of patients must fall into a "control" group.

VOL. 1 1955

TABLE V. RHEUMATIC VALVULAR HEART DISEASE

When 101 patients with congenital heart disease and 568 other patients not properly belonging to this material* are excluded, 1,031 patients remain. Of those, 788 were included in the material because of definite or uncertain rheumatic fever, and 243 because of the diagnosis of rheumatic valvular heart disease without acute rheumatic fever at the time of the *first* hospitalization. These two groups are designated "A" and "B."

	A		*B*	
Number of patients	788		243	
Definite valvular heart disease, presumably "rheumatic"	90	11%	140	58%
Possible valvular heart disease	29	4%	16	7%
Observation cases (only slightly suspected)	163	20%	41	16%
Valvular heart disease *not excluded* on account of too short observation time	399	51%	8	3%
Valvular heart disease *excluded*	107	14%	38	16%
Previous diagnosis of valvular heart disease *confirmed*	59		140	
Previous diagnosis of valvular heart disease *removed* or classified as only observation case	71		103	
First diagnosis made	32		—	

*In this table only the groups, rheumatic rever (definite or uncertain), chorea, erythema nodosum, and patients first diagnosed as having rheumatic valvular disease, are included.

In this group will also fall all patients with doubtful or uncertain findings, indicating a heart lesion, but not permitting a definite diagnosis. In our material, patients with slight but definite bulging of the left auricle at x-ray* in the absence

*Whereas a gross bulging of the left auricle is always pathologic, we probably know too little about the borderline cases. They represent a challenge to the investigator.

of pathologic physical findings formed an appreciable part of this group. Sometimes, the reverse situation—doubtful physical signs with normal x-ray—was encountered. Obviously, this group also comprises patients in whom valvular lesions were not suspected but other suspect findings were recorded, although no positive diagnosis was arrived at.

The re-evaluation of all cases in the light of later development necessitated quite a few revisions of the previous diagnosis. Generally, this consisted of a reduction of a positive diagnosis of valvular heart disease and a reclassification of the patient as only an "observation" case, as indicated in Table V.

This table also shows that about one-third of the patients with valvular heart lesions after rheumatic fever did not know of their valvular lesion. It is, furthermore, of interest to know that of 236 patients in the total material with definite valvular heart lesions, 204 were previously diagnosed, but only 131 of these were under regular medical supervision. The number of patients not cared for medically thus was 105 altogether, or 44 per cent. These data, to my mind, clearly demonstrate the importance of some kind of *regular* follow-up studies on patients with previous rheumatic fever. This is all the more important because many of these patients during their first acute phase belong to the realm of the pediatricians, but during the "innocent" phase of valvular scarring pass over into adult medicine without anybody being informed thereof.

As for the distribution of various valvular lesions in our material, they appear as in Table VI, with an age distribution as in Table VII.

TABLE VI. DISTRIBUTION OF RHEUMATIC VALVULAR HEART DISEASE ACCORDING TO TYPE AND SEX

	MITRAL VALVULAR LESIONS		COMBINED MITRAL AND AORTIC LESIONS	PURE AORTIC LESIONS	OTHER LESIONS	TOTAL
	PURE STENOSIS	REGURGITATION WITH OR WITHOUT STENOSIS				
Male	18 24.7%	9 12.3%	16 21.9%	27 37.0%	3 4.1%	73
Female	67 41.1%	40 24.5%	39 23.8%	13 8.0%	4 2.4%	163
Total	85 36.0%	49 20.8%	55 23.3%	40 16.9%	7 3.0%	236

The distributions shown in these two tables are in general accord with statistics from other places. They stress the fact that valvular lesions are much more frequent in women than in men, and this difference seems to be even more striking here than in the incidence of rheumatic fever. Mitral valvular disease is almost four times as frequent in females as in males. This proportion, which also holds true for our material on operated mitral stenosis, is more exaggerated in our material than in most other material published. Müller[8] in Oslo has

recently reported the same proportion in autopsy material. On the contrary, pure aortic lesions are much more frequent in males than in females.

TABLE VII. DISTRIBUTION OF VALVULAR OR CONGENITAL HEART LESIONS ACCORDING TO SEX AND AGE

YEARS		0-9	10-19	20-29	30-39	40-49	50-59	60-69	>70
Congenital heart disease	M	8	7	3	1	1	—	—	—
	F	15	15	6	4	6	—	—	—
Mitral stenosis, pure	M	—	—	2	3	7	3	3	—
	F	—	1	5	12	20	13	13	3
Mitral regurgitation	M	—	1	1	—	1	—	—	—
	F	—	—	1	1	—	—	—	—
Mitral stenosis and regurgitation	M	—	1	1	—	3	—	1	—
	F	—	2	1	2	6	17	10	—
Mitral stenosis and aortic regurgitation	M	—	1	1	3	2	2	1	—
	F	—	—	2	4	7	5	—	—
Aortic regurgitation, pure	M	—	2	3	6	3	2	1	—
	F	—	—	2	2	3	—	2	1
Aortic stenosis (and regurgitation)	M	—	—	1	—	4	2	2	—
	F	—	1	—	—	—	1	1	—
Other lesions	M	—	—	2	—	3	1	1	3
	F	—	1	1	2	6	10	4	1

Another interesting feature is the fairly high average age of these patients, and particularly in the mitral groups. The figures here are those of surviving patients, and the analysis of a study of dead patients may change the impression. From a special study performed on hospital patients with valvular heart disease in 1950[9] we have, however, reason to believe that the mean age at death in fatal cases will not substantially lower the mean age of the total population of this study. In comparison with data from 1928-1951 given by White,[10] our figures seem to represent a movement to even higher age groups than those arrived at in his 1951 material. Knowledge of the average expectation of life in these patients must have some bearing upon the indications for operative treatment of heart lesions.

The fifth and last point concerns a proper moment for decision regarding surgical intervention.

The selection of cases for surgical therapy of rheumatic valvular lesions is a matter of opinion[11] and will not be dealt with here. Whichever indications are applied, however—more rigid or more liberal—there is always in each indi-

VOL. 1 1955

vidual case an optimal point that should not be passed without a clear decision being made.* A recent survey in our country has shown that the incidence of mitral valvulotomies is much higher among patients living in the region of a center of thoracic surgery than in other regions. Obviously, through lack of adequate information to physicians and laymen, many patients may, thus, be deprived of a reasonable chance to improve their condition.

All attempts at preventive measures and early diagnosis are based to a certain extent upon a philosophy that human suffering is thus diminished and the expenses for medical care, for the individual or the society, are efficiently reduced. While there is little argument that this is true for acute infectious diseases, there seems to be some hesitation before the application of the same policy to chronic diseases or nontransmissible diseases without serious menace to the public in general. The follow-up study reported here was also designed as a general health examination of the persons in question. The incidence of some common findings is given in Table VIII. It is obvious that a considerable number of suspicious findings were made, and that further examinations had to be recommended for the elucidation of these findings. Among the more remarkable observations were, furthermore, three patients with unknown cancer of the breast and three patients with aneurysm of the aorta.

TABLE VIII. INCIDENCE OF SOME PATHOLOGIC FINDINGS IN THE TOTAL MATERIAL

	VALVULAR HEART DISEASE		"OBSERVATION" CASES		NO VALVULAR HEART DISEASE	
	Males	Females	Males	Females	Males	Females
Number	73	163	172	200	324	507
Percentage of patients with: A relative heart volume at x-ray exceeding 500 c.c. sq. M. body surface in males and 450 in females	63.0	76.3	19.3	18.5	6.8	7.3
Slight but definite bulging of left auricle	58.9	74.2	16.3	21.5	—	0.4
Pathologic electrocardiogram	54.8	65.9	19.2	13.0	2.8	3.6
Erythrocyte sedimentation rate (Westergren) above 15 mm./hour in males and 20 mm./hour in females	4.1	19.6	6.2	15.0	6.8	13.8
Albuminuria	20.5	15.3	12.2	7.1	10.2	9.3
Systolic blood pressure >160	12.3	7.4	0.3	4.0	2.8	5.3
Diastolic blood pressure >100	1.4	15.3	4.7	6.5	10.5	11.6
Overweight (exceeding 13 kilograms if height in cm.—100 is used as standard)	1.4	3.7	0.6	4.4	2.5	5.3

*Generally, in our opinion, this moment occurs in mitral stenosis when the patient is passing from the second to the third functional group according to the New York Heart Association classification.[11]

The cost of the examination was, actually, seven dollars per person because two of the participating physicians contributed their services to the research project. Under normal circumstances, with their salaries included, the cost would have mounted to between nine and ten dollars. The cost of every new "first diagnosis" of a valvular heart lesion—even in this selected group—was, actually, 400 dollars. If, however, all other definitely pathologic findings are included, this figure will be substantially reduced. Nevertheless, the medical economics of procedures of this and related kinds should be evaluated. Now that we have all these data at hand, will it pay a reasonable dividend to continue the investigation by some kind of a special, combined cardiac registry and clinic, in cooperation with the practicing physicians? We believe so.

SUMMARY

Rheumatic fever and rheumatic heart disease are examined from the point of view of a health-disease continuum. Strategic points for preventive measures and early diagnosis are indicated. Preliminary results from a recent follow-up study of 1,500 rheumatic fever patients in Malmö, Sweden, are presented as illustrative arguments. Some of the difficulties and open questions provoked by this study are detailed in the text.

VOL
1
195

REFERENCES

1. Padmavati, S.: Personal communication.
2. Edström, G.: Febris rheumatica, Lund: Berling, 1935.
3. Jacobbson, E.: Rheumatic Fever With Chorea Minor, Uppsala, 1946, Almquist & Wiksells.
4. Wood, P.: An Appreciation of Mitral Stenosis, Brit. M. J. **1**:1051, 1113, 1954.
5. Biörck, G., and Hall, P.: Follow-up Studies in Rheumatic Fever Patients, Acta rheumatologica. (In press.)
6. Bland, E. F., and Jones, T. D.: Rheumatic Fever and Rheumatic Heart Disease, Circulation **4**:836-844, 1951.
7. Jones, T. D.: The Diagnosis of Rheumatic Fever, J.A.M.A. **126**:481-484, 1944.
8. Müller, C.: Paper read before the Norwegian Society for Internal Medicine, 1954.
9. Biörck, G., et al.: Synpunkter på 1950 års vitiematerial vid Malmö Allmänna sjukhus, Svenska läk.-tidning **48**:2214-2224, 1951.
10. White, P. D.: Heart Disease, ed. 4, New York, 1951, The Macmillan Company.
11. Biörck, G.: When Should Mitral Stenosis Be Operated Upon? Cardiologia. (In press.)

SOME OBSERVATIONS ON THE COURSE OF TREATED SYPHILITIC PRIMARY OPTIC ATROPHY

RICHARD D. HAHN, M.D., HENRY E. ZELLMANN, M.D., HOWARD NAQUIN, M.D., ERNEST S. CROSS, JR., M.D., AND DAVID M. MARCUS, M.D.

BALTIMORE, MD.

From the Chronic Disease Division of the Medical Clinics and the Wilmer Ophthalmological Institute, The Johns Hopkins University and Hospital

(Received for publication Dec. 15, 1954.)

OL.
1
955

IN 1932, Moore[1] critically reviewed the studies on syphilitic primary optic atrophy published up to that time. Concerning the natural course of the disease it was possible to conclude that it is "so variable as to render unsafe any conclusions as to the beneficial effects of treatment in the individual patient." Once again, in 1940, Moore and Woods[2] stated that the available literature permitted of no valid conclusions as to the efficacy of treatment of any type. In 1942, Moore, Hahn, Woods, and Sloan[3] presented a detailed study of the Johns Hopkins material. The important conclusion reached at that time was that, *in a selected group of slow progressors*, malaria therapy seemed to be of value. The value of metal and subdural therapy seemed to be nil. Even this limited conclusion is subject to a number of valid criticisms:

1. Since all patients not already blind when first seen were given one or another form of treatment, the course of the untreated disease had of necessity to be estimated from anamnestic data alone. Such data are notoriously unreliable.

2. The malaria-treated patients, when analyzed for their course prior to treatment, represented an extremely selected group of slow progressors. They could not be compared at all with the untreated patients nor with those treated with metals, but only with the subdural-treated group, who likewise represented a group of slow progressors apparently similar to the malaria-treated group. The difficulty with this matching process is that even if one succeeds in matching on pertinent variables considered to be important, one may fail to match on other variables that influence the result, whose effect one does not know. For example, what part of the difference in outcome between the subdural- and malaria-treated groups was due to the deleterious effect of subdural treatment and what part was due to the beneficial effect of malaria treatment cannot be estimated.

3. The malaria-treated group was comprised of only thirty-two patients, whose course was estimated beginning with onset of symptoms rather than with

Aided by a grant-in-aid from the U. S. Public Health Service, Department of Health, Education, and Welfare.

first observation. Even including this anamnestic period the average period of "follow-up" (not observation in the true sense of the word) was only four years.

4. Blindness, defined as visual acuity of 10/200 or less in the better eye, was taken as the sole criterion of success or failure of treatment. Regardless of definition, blindness, which is an end point, is a poor criterion. Thus if 10/200 were used, a patient whose vision falls from 20/20 to 12/200 would be accounted a success and one whose vision falls from 12/200 to 10/200 a failure. This is obviously misleading. The real issue in judging the possible efficacy of treatment is its ability to prevent progression.

Since the 1942 paper, nothing approaching a statistically valid study has appeared in the literature. It will suffice to cite numbers and observation periods to prove the correctness of this statement.

McLean[4] studied 30 patients, 21 of whom had received malaria therapy. The average observation period was seven and one-half years. Levin and his associates[5] studied 54 patients, of whom only 16 had received fever therapy. Only 50 per cent of the patients were followed more than two years after treatment. Klauder,[6] in 1948, reported the results of penicillin treatment in 49 patients. Twenty-seven of these were "industrially blind" before treatment. Four patients had received fever therapy before penicillin, 17 concurrently with penicillin, and an unstated number subsequent to penicillin. An unstated number received repeated courses of penicillin. The period of observation ranged from a few months to a maximum of forty-eight months. Progression was noted in 10 cases and blindness in 6 cases. Klauder stated, "I would hesitate to draw conclusions." Also in 1948, Horne, Curtis, and Norton[7] reported on 4 patients treated with penicillin alone and 7 treated with penicillin plus malaria followed for two years after treatment. Treatment failure (undefined) occurred in none of the former and 2 of the latter group. In the same year, LaMotte, Stokes, and others[8] reported on the results of therapy with penicillin alone in 31 patients with primary optic atrophy followed from six months to four years. Ten became worse and 21 remained stable. The authors concluded, "The effectiveness of penicillin will be properly evaluated only after another 4-5 year period of follow-up." Benton and Heyman[9] reported on the penicillin treatment of 12 patients with primary optic atrophy, in 6 of whom progression to blindness or near blindness was noted. In 1949, Horne and Curtis[10] and Klauder and Gross[11] presented papers essentially identical with those presented by the same authors in 1948, cited previously.

Blatt[12] stated that he had seen 387 cases of syphilitic optic atrophy. He remarked, "It seems superfluous to me to state in detail the statistical records of the cases I followed in the course of 20 years. . . ." All of his 119 malaria-treated patients were blind within two years, a result which was worse than with metal treatment or without any treatment. This paper by Blatt, although statistically worthless, is perhaps an important illustration of the complete lack of agreement in the literature concerning the value of fever treatment.

Spiller and Stewart[13] reported on 27 patients with optic atrophy treated with penicillin with or without malaria, as part of a larger study of neurosyphilis in general. They concluded that the response in the two treatment groups was

similar. Hahn,[14] also as part of a study of neurosyphilis in general, reported progression in 6 of 17 patients with primary optic atrophy after treatment with penicillin alone, and in 3 of 16 after treatment with penicillin plus malaria. The median period of follow-up was only fifteen months. He stated, "It is impossible to draw any valid conclusions as to the comparative efficacy of penicillin and penicillin plus malaria therapy from such small numbers of patients followed for so brief a time." This conclusion might well apply to all other studies thus far available.

Epstein and Allen[15] treated 12 patients with combined penicillin and artificial fever therapy. Progression was noted in 4 of the 12. Klauder[16] reported on the results of treatment of 80 patients, 42 of whom were "industrially blind" before treatment. Most of these patients (exact number unstated) apparently were given penicillin plus malaria, although "some" were given penicillin alone. Of these 80 patients, 50 showed no change over a follow-up period of twelve to seventy-eight months, 15 were worse, and another 15 were blind. Benton and Harris[17] found progressive visual loss in 13 of 23 patients followed twenty to sixty-seven months after penicillin therapy.

In 1953, Kenney and Curtis[18] reported the results of treatment in 37 patients observed for four to eight years. Fifteen were treated with penicillin alone, and 22 with penicillin and malaria. Eighteen of the 37 patients were industrially blind in both eyes before treatment. Five of the penicillin-treated patients and 7 of the penicillin- and malaria-treated patients experienced progression of the disease after treatment. The authors concluded, "It does not appear that therapeutic malaria adds enough to the penicillin regimen to warrant its use in the treatment of early syphilitic optic atrophy."

In 1954, Klauder and Gross[19] reported the results of treatment "with penicillin" in 104 patients. This paper virtually defies analysis. No less than 50 of the 104 patients were "industrially blind" before treatment and therefore could not progress very much, if at all. Moreover, 37 of the 104 patients received fever therapy along with the first course of penicillin, and additional patients received fever therapy, metals, and cortisone, plus more penicillin as re-treatment if progression was unfavorable. Twenty-four of the 104 patients "became blind." Of 86 nonpenicillin-treated patients included for comparison, 30 "became blind." However, 38 of these 86 patients were industrially blind before treatment. The authors stated, "It is difficult to define the value if any of metal chemotherapy in this study." "It is difficult to formulate and to define the value of treatment." With this latter statement, at least, there can be no disagreement.

It seems apparent that after a lapse of over twenty years since Moore's[1] previously quoted 1932 conclusions, we are little better off than we were at that time with respect to an adequately documented, biostatistically valid demonstration of the efficacy of any form of treatment. The chief reason for the relative lack of value of the literature on the treatment of syphilitic optic atrophy is the total absence of untreated controls in attempting to evaluate progression and treatment in such a chronic variable disease. To estimate the efficacy of therapeutic procedures in any disease one should accurately know the prognosis

in the absence of treatment. In the case of syphilitic optic atrophy this requirement has not been met. Few of the patients in our previous study or in any other so far reported have been allowed to remain untreated after initial observation, and the course of untreated optic atrophy had of necessity to be estimated from anamnestic data influenced by a multiplicity of extraneous factors and derived largely from the group of patients already blind when first seen.*

Our present study is subject to the same criticism as all previous studies in that no comparable untreated group of patients has been observed. Moreover, strict definition of material has greatly reduced the numbers of cases available for analysis. It has nevertheless seemed worth while to present our data, if only as an example of the difficulties inherent in a retrospective study of a variable chronic disease process.

For the purpose of the present study and in order to ensure as great objectivity as possible, the course of the disease is considered to date from the time of the initial Johns Hopkins Hospital observation rather than from the onset of symptoms. The life-table method is utilized to describe this course between initial and latest observation. The time interval from presumed onset of symptoms to our own initial observation is utilized only in a subsidiary manner in order to estimate the course prior to our treatment. Even this portion of the course was frequently complicated by treatment which had been given elsewhere prior to our initial observation. In brief, the period under study is, in the present instance, one of true observation at our own institution.

Only patients with visual acuity greater than 10/200 at the time of our initial observation are included, thus eliminating the fallacy of considering nonprogression of an "industrially blind" patient to be a treatment success. The ophthalmologic diagnosis in all cases was established by a member of the Department of Ophthalmology of The Johns Hopkins Hospital. Either clinical or spinal fluid evidence of neurosyphilis was also a criterion for inclusion in order to ensure that syphilis and not some other condition was the cause. Excluded were all patients blind at the time of our initial observation, all patients with

*It is not implied that nothing is known of the course of the untreated disease. On the contrary, in the prearsphenamine era, bilaterality of the optic atrophy was uniform and progression to blindness or near blindness was apparently almost so. The older literature, however, made no pretense of an estimate of the average interval from onset of symptoms to blindness. Moore, Hahn, Woods, and Sloan[3] estimated the course of the untreated disease by correlating pretreatment visual acuity with duration of symptoms as related by the patient. It was found that blindness (defined as visual acuity of 10/200 or less in the better eye) occurred in 28 per cent of patients within one year of onset of symptoms, in 50 per cent within two years, in 65 per cent within three years, and in 75 per cent within six years. It is unlikely that significant alteration of these figures, based upon 252.5 person-years of retrospective analysis, will result from any future study.

Nevertheless, it is manifestly not a sound device to compare observed course after treatment with historical data concerning estimated course prior to observation and treatment. The two sets of data are not comparable. The former is objective, the latter subjective and influenced by a multiplicity of extraneous factors. Duration of symptoms is not synonymous with duration of the lesion in the optic nerve which, a priori, must have progressed to a greater or lesser degree before the patient will have noted visual failure. Other factors influencing the apparent duration of symptoms are the care with which the history is taken; the accuracy of the patient's memory and powers of observation; unilateral onset, in which one eye may become completely blind before the other is involved at all, and before the patient is aware that he has only monocular vision; and the type of visual field defect. A central scotoma will obviously precipitate subjective symptoms of visual failure earlier than will concentric constriction of the visual fields. In the analyses of individual case records it was repeatedly noted that the elicited onset of symptoms varied from examiner to examiner.

optic atrophy secondary to widespread definite inflammatory disease of the globe or with optic neuritis, all patients with acute syphilitic meningitis, and all patients with other ocular disease (e.g., cataracts) insufficiently severe to interfere with mensuration of the course of optic atrophy. Of 561 case histories reviewed, 233 were found to be suitable for analysis as to basic data. As will be detailed subsequently, this number was further considerably reduced with respect to evaluation of treatment, due to total absence of follow-up observations in many patients.

A code sheet was drawn up for each patient and the data transferred to punched cards. No patient was utilized in more than one treatment group. The percentage of visual efficiency was determined for each eye using the visual acuity (corrected for refractive error whenever available) recorded in the charts and conversion table (Appendix I) as recommended by the Report of the Committee on Compensation for Eye Injuries of the American Medical Association.[20] Visual efficiency for both eyes was determined by adding visual efficiency for each eye and dividing by two.

OL.
1
955

Since, as previously stated, the real issue in judging the possible efficacy of treatment is its ability to prevent progression rather than blindness, results are expressed in terms of *percentage loss* of pretreatment visual efficiency. Thus, if initial visual efficiency (both eyes) was 60 per cent and latest visual efficiency 35 per cent, percentage loss of visual efficiency was $\frac{60 - 35}{60}$ or 41.6 per cent. If initial visual efficiency was 90 per cent and latest visual efficiency was 35 per cent, percentage loss of visual efficiency was $\frac{90 - 35}{90}$ or 61.1 per cent, thus reflecting the greater degree of progression in the latter instance even though latest visual efficiency was the same in both cases.

Due to the relatively small numbers of patients and the often multiple treatments it was not possible to consider each form of treatment separately. Treatment groups, based solely upon treatment given *after* the initial Johns Hopkins Hospital observation, are as follows:

1. No treatment after initial Johns Hopkins Hospital observation (heterogeneous treatment elsewhere previously),
2. Metal chemotherapy (with or without subdural therapy),
3. Malaria therapy without penicillin therapy (with or without metals and subdural therapy),
4. Penicillin and malaria therapy,
5. Penicillin without malaria.

In Table I is shown treatment given elsewhere prior to our initial observation. The "no treatment" group is in fact a heterogeneous group largely treated elsewhere by various methods. The metal chemotherapy group is relatively pure in that none had previously received penicillin and only 3 per cent had previously received malaria therapy. The malaria without penicillin group, the penicillin

and malaria group, and the penicillin without malaria groups are all mutually exclusive, as prior treatment had largely been nil or with metals alone.

TABLE I. BASIC TREATMENT GROUPS BY TREATMENT GIVEN ELSEWHERE PRIOR TO FIRST J.H.H. OBSERVATION

PRIOR TREATMENT GIVEN ELSEWHERE	BASIC TREATMENT GROUPS ACCORDING TO TREATMENT GIVEN AFTER FIRST J.H.H. OBSERVATION									
	NO TREATMENT AFTER 1ST OBS.		METALS ± SUBDURAL		MALARIA WITHOUT PENICILLIN		MALARIA AND PENICILLIN		PENICILLIN WITHOUT MALARIA	
	NO.	%	NO.	%	NO.	%	NO.	%	NO.	%
None	13	31	44	63	25	40	9	28	9	33
Metals	17	40	20	28	36	58	17	53	15	55
Penicillin	5	12	—	—	—	—	5	15	1	4
Malaria	6	14	2	3	—	—	—	—	—	—
Subdural	—	—	2	3	—	—	—	—	—	—
Unknown	1	2	2	3	1	2	1	3	2	8
Total	42	100	70	100	62	100	32	100	27	100

TABLE II. BASIC TREATMENT GROUPS BY RACE AND SEX

RACE AND SEX	NO J.H.H. TREATMENT (%)	METALS (%)	MALARIA WITHOUT PENICILLIN (%)	MALARIA AND PENICILLIN (%)	PENICILLIN ALONE (%)	TOTAL (%)
W, M	37	34	60	40	18	40
N, M	38	44	24	34	68	39
W, F	13	13	13	15	3	12
N, F	12	7	3	10	11	8
White	50	47	73	55	21	52
Negro	50	51	27	44	79	47
Male	75	78	84	74	86	79
Female	25	20	16	25	14	20

Comparison of different treatment methods is best made when these methods are as nearly concurrent in time as possible. The present material falls far short of this ideal. The metal chemotherapy group were all first observed between

VOL. 1 1955

1920 and 1943. Eighty-six per cent of the malaria without penicillin group were concurrent with the metal group, the remainder being scattered through 1948. The malaria plus penicillin and the penicillin alone groups, of course, have been first observed from 1943 on. The heterogenous "treated elsewhere" group were scattered throughout the time period.

The different treatment groups are not strictly comparable with respect to age. Malaria therapy was by and large given predominantly to younger patients, only 15 per cent of each of the two malaria-treated groups being 50 years of age or older, as compared with 36 per cent of the metal-treated group, 34 per cent of the "treated elsewhere" group, and 48 per cent of the penicillin alone group.

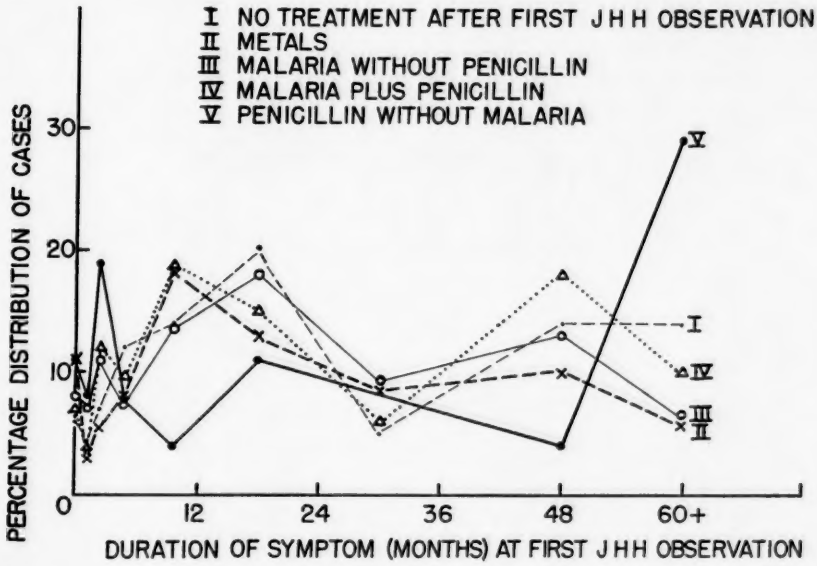


Fig. 1.—Duration of visual symptoms at first J.H.H. observation by basic treatment groups.

Percentage distributions for the total and basic treatment groups by race and sex, when known, are shown in Table II. Males predominate over females in all treatment groups. With respect to race there are striking dissimilarities between the basic treatment groups. Malaria without penicillin was given in a ratio of almost 3 to 1 to whites, penicillin without malaria was given in a ratio of over 3 to 1 to Negroes, while in the other three treatment groups, the ratio approached 1 to 1.

Fig. 1 shows duration of visual symptoms, when known, at first Johns Hopkins Hospital observation by basic treatment groups. Considering the vagaries of history taking, all treatment groups except those given penicillin without malaria are quite similar. The penicillin without malaria group contains a larger percentage of patients with symptoms of more than sixty months' duration. This may well represent merely more careful history taking in recent years, but on the other hand may mean that this group contains a larger percentage of slow progressors or of more advanced disease than do the other groups.

Fig. 2 shows percentage visual efficiency at first observation by basic treatment groups. The majority of patients in all groups had less than 50 per cent visual efficiency when first seen, and all except a very few had some loss of visual efficiency. There are no striking differences between the basic treatment groups. The penicillin without malaria group shows a slightly larger proportion of patients with lower visual efficiency than do the other treatment groups, a fact which is consistent with longer duration of symptoms in this than in the other groups.

Not shown in graphic form are scatter charts which were prepared depicting the relationship between the patient's history of duration of symptoms and visual efficiency at first observation. Suffice it to say that there was a true scatter with no direct linear relationship between these two variables. However, for any given time period after onset of symptoms, visual efficiency was approximately equally maintained in the five basic treatment groups, confirming the impressions gained from Figs. 1 and 2.

VOL 1 195

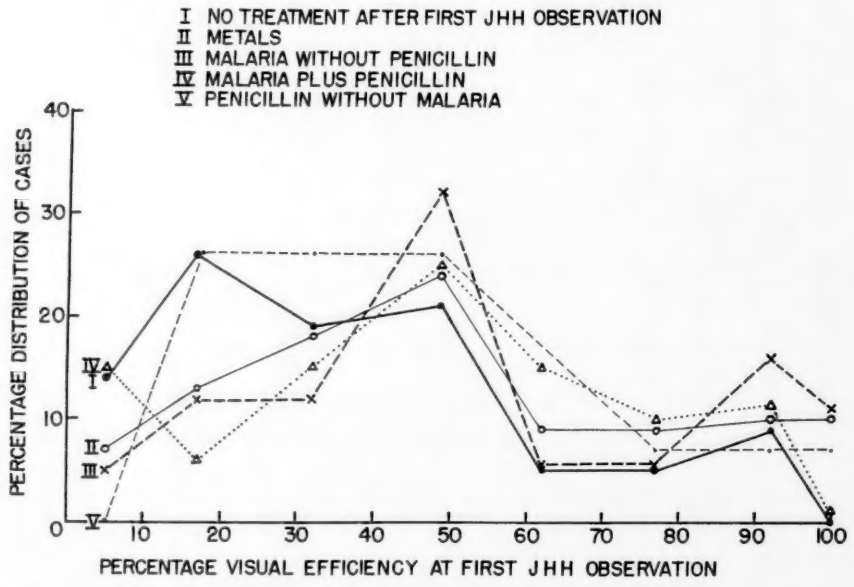


Fig. 2.—Percentage visual efficiency at first J.H.H. observation by basic treatment groups.

The inclusion of the type and degree of visual field defect in the present analysis has not proved feasible for several reasons. First, it is difficult to express such change numerically in a form suitable for comparative analysis. Second, the data available for many patients do not permit an accurate evaluation of visual field changes. A study of individual case records has revealed such variation in both peripheral and central fields, as charted by different observers, as to preclude their inclusion in an analysis based upon objective data. Thirty-two of the 233 patients had no proper charting of fields at initial observation, 12 had full fields, 32 had slightly constricted fields, 82 had moderately constricted

fields, and 75 had severely constricted fields. The recently treated patients (malaria plus penicillin and penicillin without malaria groups) had initial fields charted with but two exceptions. In other respects the basic treatment groups did not differ materially with respect to degree of constriction of initial visual fields.

The optic atrophy was considered to be unilateral in only 5 of the 233 patients. In 9 of the 233 patients, there had been some discussion among the ophthalmologists as to whether the condition was primary or secondary optic atrophy with the ultimate conclusion that it was probably primary. In 3 of the 233 patients there were peripheral retinal lesions, considered to be independent of the optic atrophy. In approximately 50 per cent of the cases the initial diagnosis was confirmed by more than one ophthalmologist. Most of the remainder ultimately were also seen by more than one ophthalmologist.

The neurological diagnosis in 72 patients was tabes dorsalis, in 19 patients taboparesis, in 18 patients paresis, in 27 patients unclassified neurosyphilis, and in 87 patients the optic atrophy with or without pupillary changes was recorded as the only significant finding. In 10 patients, neurological examination was not sufficiently complete for categorization. The basic treatment groups differed somewhat with respect to initial spinal fluid group. It will be noted (Table III) that the more recently treated patients (malaria plus penicillin and penicillin without malaria groups) had more strongly positive spinal fluids. It is believed that this is in large part a function of increased sensitivity of spinal fluid Wassermann tests in recent years, rather than a true difference in case material. Table III also shows that incomplete spinal fluid data have become less frequent in recent years.

TABLE III. BASIC TREATMENT GROUPS BY INITIAL SPINAL FLUID GROUP*

INITIAL SPINAL FLUID GROUP*	NO J.H.H. TREATMENT (%)	METALS (%)	MALARIA WITHOUT PENICILLIN (%)	MALARIA AND PENICILLIN (%)	PENICILLIN ALONE (%)	TOTAL (%)
Negative	19	14	3	3	11	11
I	—	2	2	—	—	1
II	21	27	39	40	26	31
III	30	33	48	53	60	41
Not Grouped	30	23	9	3	4	16

*Group I—Wassermann negative. Protein more than 50 mg. per cent and/or mastic 3 or more and/or cells 10 or more.
Group II—Wassermann definitely positive. Includes all such cases not in Group III.
Group III—Wassermann positive with 0.2 c.c. or less, protein more than 50 mg. per cent, mastic 4 or more in first zone.
Not grouped—Incomplete data or unknown.

In Table IV is shown initial spinal fluid cell count. It should be noted that this initial cell count was not infrequently influenced by previous treatment

given elsewhere (see Table I). All the basic treatment groups are essentially similar with respect to percentage of inactive fluids. Patients not treated after first observation and patients given metals showed a lower percentage of active fluids and a higher percentage of unknowns than did patients treated by the other methods. With respect to cell counts, just as with respect to spinal fluid groups, it is probable that these differences are spurious and due to more careful analysis of the fluids in recent than in earlier years.

TABLE IV. BASIC TREATMENT GROUPS BY INITIAL SPINAL FLUID CELL COUNT

INITIAL SPINAL FLUID CELL COUNT	NO J.H.H. TREATMENT (%)	METALS (%)	MALARIA WITHOUT PENICILLIN (%)	MALARIA AND PENICILLIN (%)	PENICILLIN ALONE (%)	TOTAL (%)
Inactive (0-5)	35	28	31	37	33	32
Borderline (6-10)	5	7	2	13	8	6
Active (11 or more)	29	32	49	47	51	38
Unknown	31	33	18	3	8	24

TABLE V. NUMBER AND PERCENTAGE OF PATIENTS IN EACH BASIC TREATMENT GROUP WITH SECOND OR SUBSEQUENT J.H.H. OPHTHALMOLOGIC OBSERVATIONS

TREATMENT GROUP	TOTAL NUMBER	NUMBER WITH TWO OR MORE OBSERVATIONS	PER CENT WITH TWO OR MORE OBSERVATIONS
No treatment after first observation	42	17	40
Metals ± subdural	70	39	56
Malaria without penicillin	62	48	77
Malaria plus penicillin	32	30	94
Penicillin without malaria	27	24	89

In Table V are shown the number and per cent of patients in each basic treatment group for whom one or more follow-up ophthalmologic observations at this institution were available. Follow-up of the "treated elsewhere—no treatment at Johns Hopkins Hospital" group was so poor as to preclude any attempt at determination of outcome for this particular group. Treatment for this group had been so heterogeneous that no conclusions could have been drawn under any circumstances. The percentage of patients followed was poor for the metal-treated group and relatively good for the more recently treated malaria, malaria plus penicillin, and penicillin groups. This difference in excellence of follow-up introduces yet another serious bias in comparison between

treatment groups, particularly between those treated with metals and those given penicillin, malaria, or both. Since only 56 per cent of the metal-treated patients returned for a second observation, any such comparison necessitates the large assumption that the 44 per cent who were not followed were similar to the 56 per cent who were followed.

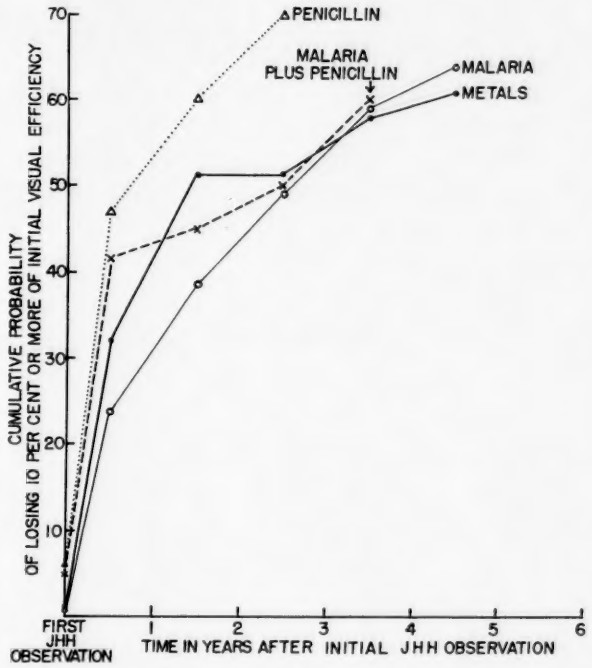


Fig. 3.—Cumulative probability of losing 10 per cent or more of initial visual efficiency by basic treatment groups.

In Table VI and Fig. 3 are depicted the cumulative probability of losing 10 per cent or more of initial visual efficiency according to type of treatment received. The actual life tables upon which subsequent figures are based have been omitted, but in general, life-table calculations were terminated when only one-third of patients remained under observation. In consideration of the numerous previously discussed uncontrolled variables, and the small numbers of cases, differences between the basic treatment groups are not considered to be significant and certainly not striking. The one conclusion which it seems permissible to draw from Fig. 3 is that no form of treatment seems capable of completely arresting progression in a high proportion of cases. Thus at two to three years after initial observation, 51.5 per cent of the metal-treated group, 49 per cent of the malaria-treated group, 49.3 per cent of the malaria-plus-penicillin—treated group, and 70.3 per cent of the penicillin group had lost 10 per cent or more of initial visual efficiency.

TABLE VI. CUMULATIVE PROBABILITY OF LOSING 10 PER CENT OR MORE OF INITIAL VISUAL EFFICIENCY BY BASIC TREATMENT GROUPS

TREATMENT GROUP	INTERVAL (YR. AFTER 1ST OBSERVATION)	O_x (NO. UNDER OBSERVATION WITHOUT 10% LOSS AT BEGINNING OF INTERVAL)	d_x (NO. WITH 10% LOSS DURING THIS INTERVAL)	W_x (NO. LAST SEEN DURING THIS INTERVAL WITHOUT 10% LOSS)	Q_x (PROBABILITY OF 10% LOSS DURING THIS INTERVAL) $\left(\frac{d_x}{O_x - \frac{W_x}{2}}\right)$	P_x (PROBABILITY OF NOT LOSING 10% DURING THIS INTERVAL) $(100 - Q_x)$	CP_x (CUMULATIVE PROBABILITY OF NOT LOSING 10%) $(P_{x_n} \cdot P_{x_{n+1}})$	CQ_x (CUMULATIVE PROBABILITY OF LOSING 10%) $(100 - CP_x)$
Metals	<1	39	12	2	31.6	68.4	68.4	31.6
	1–2	25	7	2	29.1	70.9	48.5	51.5
	2–3	16	0	0	0	100.0	48.5	51.5
	3–4	16	2	0	12.5	87.5	42.4	57.6
	4–5	14	1	1	7.4	92.6	39.3	60.7
Malaria	<1	48	11	5	24.2	75.8	75.8	24.2
	1–2	32	6	2	19.4	80.6	61.1	38.9
	2–3	24	4	0	16.6	83.4	51.0	49.0
	3–4	20	4	1	20.5	79.5	40.5	59.5
	4–5	15	2	0	13.3	86.7	35.1	64.9
Marlaria + Penicillin	<1	30	12	2	41.4	58.6	58.6	41.4
	1–2	16	1	1	6.5	93.5	54.8	45.2
	2–3	14	1	1	7.4	92.6	50.7	49.3
	3–4	12	2	5	21.1	78.9	40.0	60.0
Penicillin	<1	24	11	1	46.8	53.2	53.2	46.8
	1–2	12	3	1	26.1	73.9	39.3	60.7
	2–3	8	2	0	25.0	75.0	29.5	70.5

VOL 1 195

In Fig. 4 is depicted the cumulative probability of losing 80 per cent or more of initial visual efficiency, according to type of treatment received. Once again differences between treatment groups do not appear to be significant. Fig. 4 is, however, of considerable interest in that it gives an entirely different picture of the prognosis of syphilitic optic atrophy under direct observation as compared with that derived from questioning already blind patients as to the duration of their symptoms. At three to four years after our first observation only 17.1 per cent of the metal-treated group, 5.5 per cent of the malaria-treated group, 11.8 per cent of the malaria-plus-penicillin–treated group, and 23.5 per cent of the penicillin-treated group had lost 80 per cent or more of their initial visual efficiency. Even at ten to eleven years only 21.3 per cent and 13 per cent of the metal and malaria groups, respectively, had sustained a similar degree of loss.

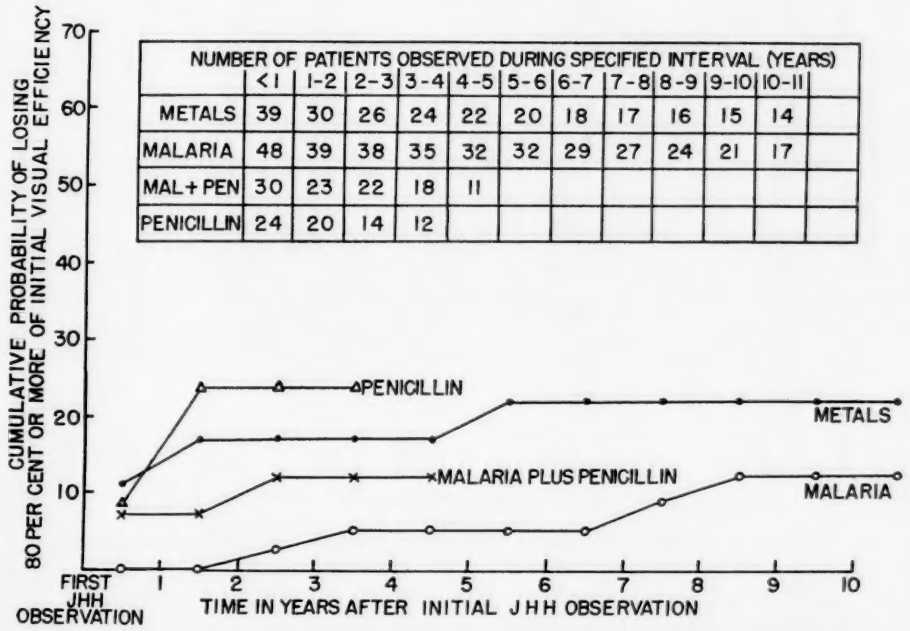
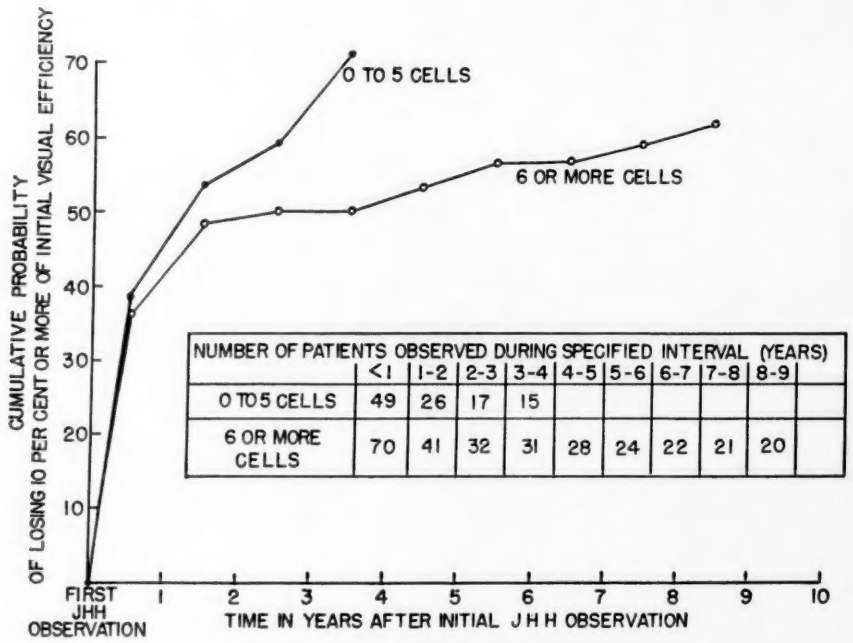

NUMBER OF PATIENTS OBSERVED DURING SPECIFIED INTERVAL (YEARS)
<1 | 1-2 | 2-3 | 3-4 | 4-5 | 5-6 | 6-7 | 7-8 | 8-9 | 9-10 | 10-11
METALS 39 30 26 24 22 20 18 17 16 15 14
MALARIA 48 39 38 35 32 32 29 27 24 21 17
MAL+PEN 30 23 22 18 11
PENICILLIN 24 20 14 12
CUMULATIVE PROBABILITY OF LOSING 80 PER CENT OR MORE OF INITIAL VISUAL EFFICIENCY
70
60
50
40
30
20
10
PENICILLIN
METALS
MALARIA PLUS PENICILLIN
MALARIA
FIRST JHH OBSERVATION
1 2 3 4 5 6 7 8 9 10
TIME IN YEARS AFTER INITIAL J H H OBSERVATION

Fig. 4.—Cumulative probability of losing 80 per cent or more of initial visual efficiency by basic treatment groups.

The variable course of syphilitic primary optic atrophy suggests the possible influence of factors other than type of treatment administered in determining outcome. In general, neurosyphilis with an active spinal fluid responds better to antisyphilitic treatment than does neurosyphilis with an inactive spinal fluid. In order to determine the possible effect of pretreatment spinal fluid cell count, patients with primary optic atrophy were subdivided into two groups irrespective of type of treatment administered. Two or more observations were available for 49 patients with initial spinal fluid cell count of less than 6 cells, and for 70 patients with initial spinal fluid cell count of 6 or more cells.

In Fig. 5 is depicted the cumulative probability of losing 10 per cent or more of initial visual efficiency according to initial spinal fluid cell count. The probability of progression is somewhat less in the presence of an active pretreatment spinal fluid. Thus at three to four years after initial observation, 50 per cent of patients with an active fluid as compared with 72.8 per cent of patients with an inactive fluid had lost 10 per cent or more of initial visual efficiency. Even at eight to nine years, only 61.5 per cent of patients with an active fluid showed similar progression.



NUMBER OF PATIENTS OBSERVED DURING SPECIFIED INTERVAL (YEARS)										
	<1	1-2	2-3	3-4	4-5	5-6	6-7	7-8	8-9	
0 TO 5 CELLS	49	26	17	15						
6 OR MORE CELLS	70	41	32	31	28	24	22	21	20	

Fig. 5.—Cumulative probability of losing 10 per cent or more of initial visual efficiency by initial spinal fluid cell count.

Apparently somewhat paradoxically but actually well within sampling variation, as depicted in Fig. 6, the cumulative probability of losing 80 per cent or more of initial visual efficiency is slightly less in the presence of an inactive than of an active pretreatment spinal fluid, being 4.7 per cent as compared with 11.7 per cent at four to five years, and 4.7 per cent as compared with 17.6 per cent at ten to eleven years after initial observation.

Fig. 6, like Fig. 5, shows a relatively favorable prognosis under direct observation for retention of some vision among those patients who are not "industrially blind" when first observed. Long-continued low-grade visual efficiency after a relatively rapid initial downhill course has long been a well-recognized clinical phenomenon in syphilitic primary optic atrophy under direct observation. This affords a possible explanation for the slightly better showing in this regard of patients with an inactive pretreatment fluid, in some of whom the process was

probably "burnt-out" before treatment was given, either spontaneously or as a result of treatment given elsewhere prior to our initial observation.

In order to evaluate the more or less generally held clinical impression that treatment administered while vision is still relatively good is more efficacious than is treatment given after vision has greatly deteriorated, patients were subdivided into two groups irrespective of type of treatment administered. Two or more observations were available for 93 patients with initial visual efficiency of 40 per cent or more, and for 65 patients with initial visual efficiency of less than 40 per cent.

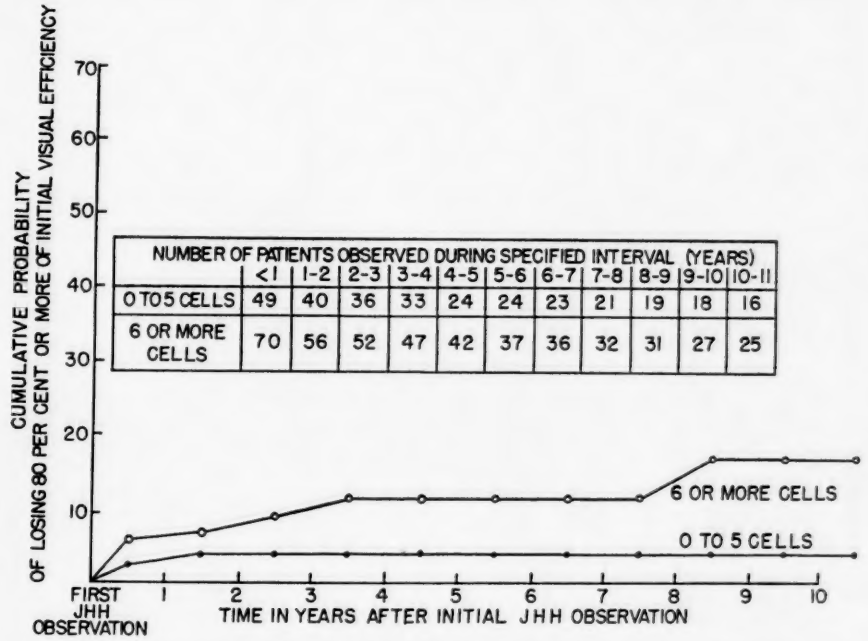

CUMULATIVE PROBABILITY OF LOSING 80 PER CENT OR MORE OF INITIAL VISUAL EFFICIENCY
70
60
50
40
30
20
10
NUMBER OF PATIENTS OBSERVED DURING SPECIFIED INTERVAL (YEARS)
<1 | 1-2 | 2-3 | 3-4 | 4-5 | 5-6 | 6-7 | 7-8 | 8-9 | 9-10 | 10-11
0 TO 5 CELLS | 49 | 40 | 36 | 33 | 24 | 24 | 23 | 21 | 19 | 18 | 16
6 OR MORE CELLS | 70 | 56 | 52 | 47 | 42 | 37 | 36 | 32 | 31 | 27 | 25
6 OR MORE CELLS
0 TO 5 CELLS
FIRST JHH OBSERVATION
1 2 3 4 5 6 7 8 9 10
TIME IN YEARS AFTER INITIAL JHH OBSERVATION

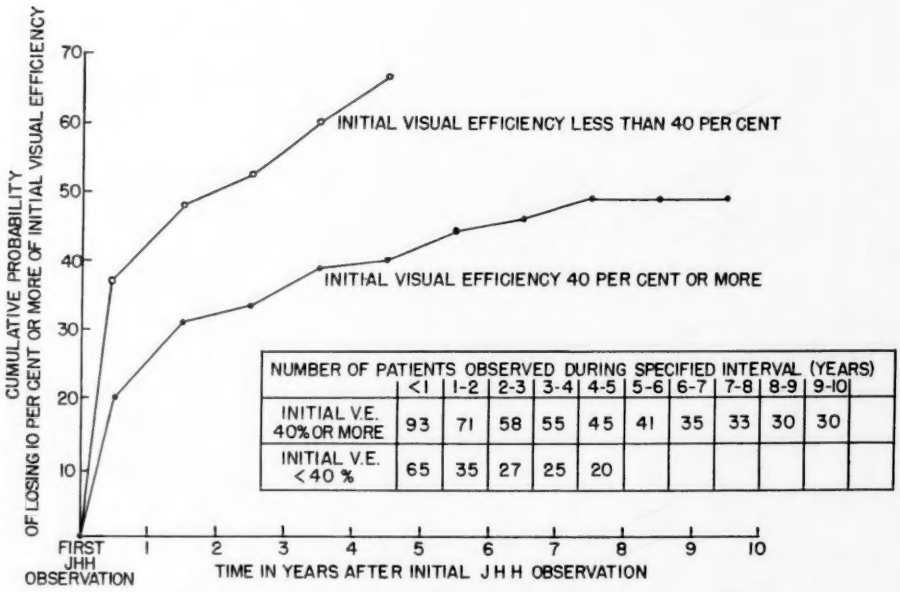
Fig. 6.—Cumulative probability of losing 80 per cent or more of initial visual efficiency by initial spinal fluid cell count.

In Fig. 7 is depicted the cumulative probability of losing 10 per cent or more of initial visual efficiency according to initial visual efficiency. The probability of progression is significantly less when treatment is given before initial visual efficiency has dropped below 40 per cent. Thus at four to five years after initial observation, 40.2 per cent of patients with initial visual efficiency of 40 per cent or more as compared with 66 per cent of patients with initial visual efficiency of less than 40 per cent had lost 10 per cent or more of initial visual efficiency. Even at nine to ten years, only 49.6 per cent of patients with initial visual efficiency of 40 per cent or more showed similar progression.

In Fig. 8 is shown the cumulative probability of losing 80 per cent or more of initial visual efficiency according to initial visual efficiency. Fig. 8 is entirely confirmatory of Fig. 7. At four to five years after initial observation only 4.8

per cent of patients with initial visual efficiency of 40 per cent or more had lost 80 per cent or more of initial visual efficiency, as compared with 14.6 per cent of patients with initial visual efficiency of less than 40 per cent. At ten to eleven years only 8.9 per cent of the former group as compared with 25.8 per cent of the latter group showed similar progression. Even at fifteen to sixteen years only 11.6 per cent of patients with initial visual efficiency of more than 40 per cent had lost more than 80 per cent of this visual efficiency.

Fig. 8 also affords rather striking confirmation of Figs. 4 and 6 in that it also shows a relatively favorable prognosis, under direct observation, for retention of some vision among those patients who are not "industrially blind" when first observed.



NUMBER OF PATIENTS OBSERVED DURING SPECIFIED INTERVAL (YEARS)	<1	1-2	2-3	3-4	4-5	5-6	6-7	7-8	8-9	9-10	
INITIAL V.E. 40% OR MORE	93	71	58	55	45	41	35	33	30	30	
INITIAL V.E. <40 %	65	35	27	25	20						

Fig. 7.—Cumulative probability of losing 10 per cent or more of initial visual efficiency by initial visual efficiency.

In Fig. 9 is shown the cumulative probability of losing 10 per cent or more and 80 per cent or more of initial visual efficiency in 141 patients after various types of antisyphilitic treatment. The data of Fig. 9 are highly reliable because they are based upon the direct observation of a relatively large group of patients. They are considered accurately to portray the course of treated syphilitic primary optic atrophy. With respect to loss of 10 per cent or more of initial visual efficiency, there is a sharp initial rise to 33.8 per cent during the first year of observation, indicating that approximately one-third of the patients are rapid progressors. Thereafter, the slope of the curve is more gradual, with approximately one-half the patients progressing by two to three years and approximately two-thirds progressing by four to five years.

With respect to loss of 80 per cent or more of visual efficiency, the prognosis is good. The slope of the curve is gradual from the onset and increases slowly and at a relatively uniform rate throughout the observation period. By one to two years only 10 per cent of patients have attained this degree of progression. By four to five years only 15 per cent and by seven to eight years only 20 per cent have shown similar progression.

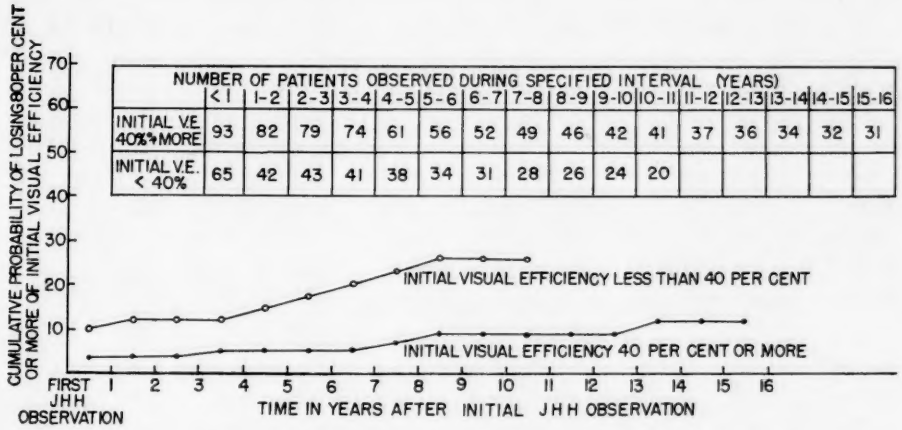

CUMULATIVE PROBABILITY OF LOSING 80 PER CENT OR MORE OF INITIAL VISUAL EFFICIENCY
70
60
50
40
30
20
10
NUMBER OF PATIENTS OBSERVED DURING SPECIFIED INTERVAL (YEARS)
< 1 | 1-2 | 2-3 | 3-4 | 4-5 | 5-6 | 6-7 | 7-8 | 8-9 | 9-10 | 10-11 | 11-12 | 12-13 | 13-14 | 14-15 | 15-16
INITIAL V.E. 40% & MORE: 93 | 82 | 79 | 74 | 61 | 56 | 52 | 49 | 46 | 42 | 41 | 37 | 36 | 34 | 32 | 31
INITIAL V.E. < 40%: 65 | 42 | 43 | 41 | 38 | 34 | 31 | 28 | 26 | 24 | 20
INITIAL VISUAL EFFICIENCY LESS THAN 40 PER CENT
INITIAL VISUAL EFFICIENCY 40 PER CENT OR MORE
FIRST JHH OBSERVATION
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16
TIME IN YEARS AFTER INITIAL JHH OBSERVATION

Fig. 8.—Cumulative probability of losing 80 per cent or more of initial visual efficiency by initial visual efficiency.

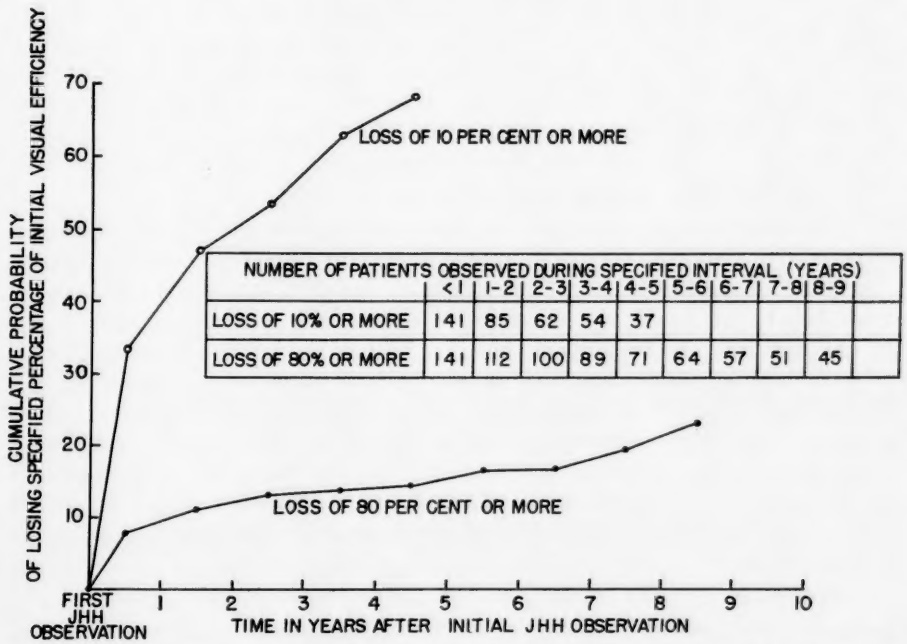

CUMULATIVE PROBABILITY OF LOSING SPECIFIED PERCENTAGE OF INITIAL VISUAL EFFICIENCY
70
60
50
40
30
20
10
LOSS OF 10 PER CENT OR MORE
NUMBER OF PATIENTS OBSERVED DURING SPECIFIED INTERVAL (YEARS)
< 1 | 1-2 | 2-3 | 3-4 | 4-5 | 5-6 | 6-7 | 7-8 | 8-9
LOSS OF 10% OR MORE: 141 | 85 | 62 | 54 | 37
LOSS OF 80% OR MORE: 141 | 112 | 100 | 89 | 71 | 64 | 57 | 51 | 45
LOSS OF 80 PER CENT OR MORE
FIRST JHH OBSERVATION
1 2 3 4 5 6 7 8 9 10
TIME IN YEARS AFTER INITIAL JHH OBSERVATION

Fig. 9.—Cumulative probability of losing specified percentage of initial visual efficiency in 141 patients after various types of treatment.

SUMMARY

A critical review of the literature, including detailed analysis of our own 1942 data, has failed to reveal a biostatistically valid study of the effect of treatment of syphilitic primary optic atrophy. To estimate the efficacy of treatment, one should accurately know the prognosis in the absence of treatment. In the case of syphilitic optic atrophy this logical requirement has not been met.

The present study is concerned with that portion of the course of the disease between our own initial and latest observations, which is subjected to analysis by the life-table method. The diagnosis both as to optic atrophy and neurosyphilis has been established with great care. Only patients with initial visual acuity greater than 10/200 are included, thus eliminating the fallacy of considering nonprogression of an "industrially blind" patient a treatment success. Visual efficiency before and after treatment was calculated from visual acuity utilizing standard conversion tables. Since prevention of progression rather than blindness is the best measure of therapy, results are expressed in terms of percentage loss of pretreatment visual efficiency.

Comparison is made between the course of the disease after metal chemotherapy (with or without subdural therapy), malaria without penicillin (with or without metals and subdural treatment), penicillin plus malaria, and penicillin without malaria. The group treated with penicillin alone contains a larger percentage of patients with symptoms of more than sixty months' duration and a slightly larger percentage with lower visual efficiency than do the other groups. However, scatter charts suggest that, in all treatment groups, pretreatment visual efficiency is approximately equally maintained for any given time period after onset of symptoms. The four treatment groups do not differ materially with respect to neurological diagnosis or degree of pretreatment visual field constriction. Differences in pretreatment spinal fluid positivity and activity are believed to be spurious and due to increased sensitivity of tests and more careful analysis in recent years.

The present study is subject to the same criticism as all previous studies in that no comparable untreated group of patients has been observed. Moreover, strict definition of material has reduced the number of cases available for analysis. Results of our treatment may have been influenced to a variable degree by previous treatment given elsewhere. The different treatment methods have been utilized at widely separated points in time. The percentage of patients followed was poor for the metal-treated group and relatively good for the other groups. Patients given malaria were younger and patients given penicillin were older than patients given metals. Malaria without penicillin was given in a ratio of almost 3 to 1 to whites, penicillin without malaria was given in a ratio of over 3 to 1 to Negroes, while in the other treatment groups the ratio approached 1 to 1.

Progression, defined as loss of 10 per cent or more of initial visual efficiency, had occurred at two to three years after initial observation in approximately 70 per cent of the penicillin-treated group and in approximately 50 per cent of the other groups. Likewise, loss of 80 per cent or more of initial visual efficiency

VOL
1
195

occurred, at three to four years, in 23.5 per cent of the penicillin group, as compared with 17.1 per cent of the metal-treated group, 5.5 per cent of the malaria-treated group, and 11.8 per cent of the malaria-plus-penicillin–treated group. In consideration of the numerous previously discussed uncontrolled variables, these differences are not considered to be significant. The somewhat poorer showing of the penicillin group is explained by lower pretreatment visual efficiency in this group.

At three to four years after initial observation, 50 per cent of patients with active fluids as compared with 73 per cent of patients with inactive fluids had progressed. Apparently somewhat paradoxically, but actually well within sampling variation, the cumulative probability of losing 80 per cent or more of initial visual efficiency is slightly less in the presence of an inactive than of an active pretreatment spinal fluid.

OL.
1
955

Patients who had 40 per cent or more visual efficiency when treated did significantly better than patients with visual efficiency of less than 40 per cent, progression occurring at four to five years in 40 per cent of the former, and 66 per cent of the latter group. Likewise, loss of 80 per cent or more of initial visual efficiency, even as long as ten to eleven years after initial observation, occurred in only 9 per cent as compared with 26 per cent.

Among 141 patients observed after various types of antisyphilitic treatment, progression occurred in approximately one-third during the first year, in approximately one-half at two to three years, and in approximately two-thirds at four to five years. Loss of 80 per cent or more of initial visual efficiency occurred in only 10 per cent at one to two years, 15 per cent at four to five years, and 20 per cent at seven to eight years.

CONCLUSIONS

1. The present study, like all previous studies, does not permit of valid comparison between the effects of the several types of treatment in syphilitic primary optic atrophy. Deficiencies in basic data preclude accurate comparison between the effects of the several treatment methods.

2. The rate of progression after treatment was not apparently significantly influenced by pretreatment spinal fluid cell count.

3. Patients treated with pretreatment visual efficiency of more than 40 per cent showed a significantly lesser rate of progression than did patients treated with initial visual efficiency of 40 per cent or less. This difference can, in our opinion, be attributed largely to the treatment given. However, even in the group with pretreatment visual efficiency ranging from 3.3 to 39 per cent, the cumulative probability of further loss of 80 per cent of pretreatment efficiency was only 26 per cent at the end of eleven years.

4. Among 141 patients observed after various types of antisyphilitic treatment, progression occurred in approximately one-third during the first year, one-half at two to three years, and two-thirds at four to five years. Loss of 80 per cent or more of initial visual efficiency occurred in only 10 per cent at one to two years, 15 per cent at four to five years, and 20 per cent at seven to eight years.

5. Perhaps the most significant conclusion to be drawn from the present study is that retrospective comparison of the effects of various treatments given during the course of long-term chronic disease, especially in the absence of an untreated control group similarly observed and similar with respect to all variables other than type of treatment received, is fraught with pitfalls for the unwary. Even if one succeeds in matching on pertinent variables considered to be

Appendix I. Table for Conversion of Visual Acuity to Visual Efficiency

VISUAL ACUITY	PERCENTAGE OF VISUAL EFFICIENCY
20/20 or better	100
20/20 or 20/30 +	95
20/30	91.5
20/30 or 20/40 +	87.5
20/40	83.6
20/40 or 20/50 +	80.0
20/50	76.5
20/50 or 20/70 +	70.0
20/70	64.0
20/70 − 1 or 20/100 + 2	58.5
20/70 − 2 or 20/100 + 1	53.4
20/100 or 40/200	48.9
19/100 or 38/200	47.0
18/100 or 36/200	44.0
17/100 or 34/200	41.0
16/100 or 32/200	38.0
15/100 or 30/200	35.0
20/100 − 1 or 20/200 + 1	35.0
14/100 or 28/200	32.0
13/100 or 26/200	29.0
12/100 or 24/200	26.0
11/100 or 22/200	23.0
10/100 or 20/200	20.0
19/200	18.0
9/100 or 18/200	16.0
17/200	14.5
8/100 or 16/200	13.0
15/200	11.5
7/100 or 14/200	9.5
13/200	8.0
6/100 or 12/200	6.0
11/200	4.5
5/100 or 10/200	3.3
9/200	2.5
4/100 or 8/200	1.4
7/200	1.0
3/100 or 6/200	0.4
5/200	0.1
Under 5/200	0.0

important, one may fail to match on other variables that influence the result, whose effect one does not know. Despite the limited conclusions drawn, it has seemed worth while to present our data in detail as an example of the difficulties inherent in a retrospective study of a variable chronic disease process.

REFERENCES

1. Moore, J. E.: The Syphilitic Optic Atrophies, With Especial Reference to Primary Optic Atrophy, Medicine **11**:263, 1932.
2. Moore, J. E., and Woods, A. C.: Syphilitic Primary Optic Atrophy. II. General Considerations and the Results of Treatment by Standard Methods (Especially Subdural Treatment and Induced Fever). A Critical Review, Am. J. Ophth. **23**:145, 1940.
3. Moore, J. E., Hahn, R. D., Woods, A. C., and Sloan, L.: The Treatment of Syphilitic Primary Optic Atrophy, Am. J. Syph., Gonor. & Ven. Dis. **26**:407, 1942.
4. McLean, J. A.: Visual Field Changes in Syphilis of Central Nervous System, Canad. M.A.J. **55**:571, 1946.
5. Levin, S., Trevett, L. D., and Greenblatt, M.: Syphilitic Primary Optic Atrophy, Review of 54 cases, New England J. Med. **237**:769, 1947.
6. Klauder, J. V.: Penicillin in the Treatment of Primary Optic Atrophy, Symposium: Recent Advances in the Study of Venereal Diseases, Syphilis Study Section, National Institutes of Health, Washington, D. C., April 8-9, 1948, p. 233.
7. Horne, S. F., Curtis, A. C., and Norton, D. H.: The Treatment of Primary Optic Atrophy With Penicillin and Malaria and Penicillin, ibid., p. 236.
8. LaMotte, W. O., Stokes, J. H., et al.: Sixty Cases of Primary Optic Atrophy Critically Studied From the Standpoint of Diagnosis and Therapy, ibid., p. 238.
9. Benton, C. D., and Heyman, A.: Treatment of Ocular Syphilis With Penicillin, Arch. Ophth. **40**:302, 1948.
10. Horne, S. F., and Curtis, A. C.: The Treatment of Syphilitic Primary Optic Atrophy With Penicillin and With Penicillin and Malaria, Am. J. Syph., Gonor. & Ven. Dis. **33**:143, 1949.
11. Klauder, J. V., and Gross, B. A.: Penicillin Treatment of Syphilitic Primary Optic Atrophy, Am. J. Syph., Gonor. & Ven. Dis. **33**:234, 1949.
12. Blatt, N.: La valeur de la malaria-thérapie dans l'atrophie syphilitique des nerfs optiques, Ann. ocul. **182**:513, 1949.
13. Spiller, W. F., and Stewart, J.: Neurosyphilis: Comparative Treatment With Penicillin Alone and Penicillin Plus Malaria, J. Invest. Dermat. **14**:121, 1950.
14. Hahn, R. D.: The Treatment of Neurosyphilis With Penicillin and With Penicillin Plus Malaria, Am. J. Syph., Gonor. & Ven. Dis. **35**:433, 1951.
15. Epstein, N. N., and Allen, J. R.: Treatment of Neurosyphilis With Penicillin Combined With Artificial Fever Therapy, Arch. Dermat. & Syph. **63**:419, 1951.
16. Klauder, J. V.: Blindness Due to Syphilis, J. Ven. Dis. Inform. **32**:183, 1951.
17. Benton, C. D., Jr., and Harris, J. F.: Syphilitic Optic Nerve Atrophy Treated With Penicillin. Observations Two to Six Years After Treatment, Arch. Ophth. **48**:449, 1952.
18. Kenney, J. A., Jr., and Curtis, A. C.: Treatment of Syphilitic Optic Atrophy by Penicillin, With and Without Therapeutic Malaria, Am. J. Syph., Gonor. & Ven. Dis. **37**:449, 1953.
19. Klauder, J. V., and Gross, B. A.: Results of Penicillin, Cortisone, and Non-penicillin Treatment of Syphilitic Optic Atrophy, With Report of Clinical Observations, Am. J. Syph., Gonor. & Ven. Dis. **38**:270, 1954.
20. Snell, A. C.: A Treatise of Medicolegal Ophthalmology, St. Louis, 1940, The C. V. Mosby Company.

OL.
1
955

THE PHYSICIAN'S RESPONSIBILITY IN TUBERCULOSIS CONTROL*

HENRY STUART WILLIS, M.D.**

CHAPEL HILL, N. C.

(Received for publication Jan. 10, 1955.)

VOL
1
195

IN DISCUSSING the physician's responsibility in tuberculosis control, it seems appropriate to comment first on the popular idea that tuberculosis needs but little controlling because it is fast tumbling under the current impact of near-sufficient beds, multiple choice of drugs, and wide application of surgical maneuvers. It is true that the incidence of mortality is declining precipitately, but morbidity rates remain relatively stationary or drop with discouraging slowness. Does the public which thinks tuberculosis is licked realize that in 1953 it caused 20,100 deaths in this country and that upward of 100,000 new cases of the disease were reported?

If only one-half of the new cases are active, the disease stands out as a bold challenge to the community's best medical, sociologic, economic, and preventive resources. This formidable challenge looks economically burdensome when we realize that each case of tuberculosis has been conservatively reckoned as costing the community $14,000 in direct expense, maintenance of dependents, and lost earning power.

1. *Responsibility of the General Practitioner.*—With this as the situation, then, what is the physician's responsibility in tuberculosis control? And which physicians share in it? The physician at the grass roots—the man in general practice—is obviously in a strategic position to suspect the disease on the basis of early symptoms and to detect it in early stages, even though many patients present few or no early symptoms. He may follow through on contacts, and may advise the public that x-ray of the chest is of the greatest importance, both in survey groups and in the individual person. He ought to suspect the disease and deliberately seek it in most likely groups and situations of economic stress and to rule it out in certain susceptible people as, for instance, those with diabetes mellitus. He needs to realize that, as the incidence of tuberculosis has declined, the tuberculin test has become a tool of increasing value to him. These make up the challenge to the family physician who could, if he would, become foremost in rendering this personal and public service.

---

*Presented at the annual meeting of the California Tuberculosis and Health Association, San Francisco, April 1, 1954.

**Superintendent-Medical Director, North Carolina Sanatoriums, Chapel Hill, North Carolina.

The general practitioner actually comes in contact with the majority of the cases of tuberculosis, for studies indicate that 70 per cent of people covered in mass surveys give the name of their private physician as the one to whom reports of the examination should be sent—another reason why tuberculosis should be a subject of live and timely interest to him, and still a challenge to his clinical skill, judgment, professional alertness, and medical competence. This is why the family physician stands in his proverbial key position. The case is referred back to him and is his patient; he is still and always the family physician and usually the friend, counselor, father-confessor whose advice is often, if not always, followed. He has a real opportunity to guide and counsel many a patient and, thus, to take part in an important health movement.

As the disease grows less prevalent, the family physician sees progressively fewer cases of pulmonary tuberculosis. He naturally pays less attention to the affections which plague him less frequently. He, therefore, sometimes exhibits a low index of suspicion for tuberculosis and thus may allow the patient to reach the latter stages of the disease because of failure to search the sputum for tubercle bacilli or seek the aid of the radiologist. The following two brief case reports illustrate the point and indicate also that the hospital staff, as well as the general practitioner, can readily make this error.

OL.
1
955

CASE 1.—I. M., a white woman aged 68, had had an earlier hospital experience with tuberculosis which had become inactive. At home for nearly two years, working around the house and performing the lighter tasks associated with housekeeping, she had done reasonably well. She had "become lost" to her physician and the Health Department when she developed fever, productive cough with occasional blood streaking, and night sweats, and was diagnosed by the family physician as having pneumonia. These symptoms cleared partially in the course of one month but residual weakness, anorexia, and persistent cough remained. The weakness became increasingly prominent before suspicion of tuberculosis prompted x-ray examination and sputum study, both of which promptly yielded the diagnosis of pulmonary tuberculosis, far advanced.

CASE 2.—G. J., a Negro man aged 34, developed fever, thoracic pain, and shortness of breath and was admitted to a local hospital, where 520 c.c. of greenish yellow fluid was aspirated. This yielded hemolytic streptococci. Tubercle bacilli were not seen. The patient improved and was discharged five days after admission. It was suggested to the patient that x-ray follow-up be done to rule out tuberculosis, but this recommendation was not arranged for. Final diagnosis was pleural effusion, left side, cause undetermined. The patient returned to his home in the belief that he had "pneumonia" which was cured—this despite the advice of a subsequent roentgenogram. Ten months later he had fulminant pulmonary tuberculosis. Apparently it was forgotten that "idiopathic" pleural effusions usually are tuberculous, even though secondary infection may be present.

Would suspicion of tuberculosis at the time of the acute symptoms have caused the physician to carry out the simple tests for its detection—in the thought that the "pneumonia" had possibly intervened in the course of the more leisurely tuberculosis or was in itself a manifestation of that disease? A considerable number of tuberculous patients present a history of acute respiratory disease, diagnosed pneumonia, prior to the final diagnosis. In each of these cases failure to suspect the disease and thus failure to carry out the simple tests for its detection allowed its progression.

As the newer trend prevails of hospitalizing the tuberculosis patient for a shorter time than was formerly the case, the general practitioner will be involved

much more in the active treatment of this disease, for which he has heretofore declined responsibility and which, until now, he may have dodged. Whether the patient may safely return home some weeks or months earlier on the average than was allowable in the prechemotherapeutic era is an unsettled question. Formerly medical men treating the disease did the best they could with the best therapy at hand and taught the patient to live with the residual of his disease. Has the new look—through drugs, surgery, and rest—brought us to the point where the disease can actually be eliminated from the body, as some think? If this is true, it would seem sound to combine all the useful routines and therapies in an effort to produce eradication, which would certainly include adequate and, probably, long stay in the sanatorium.

Whether the patient is discharged earlier or not, most often it is recommended that he continue use of drugs. Under these circumstances the general practitioner simply must learn more about tuberculosis and more about the use of the drugs. He thus establishes a sort of entente or partnership with the sanatorium personnel, to the mutual advantage of both and to the benefit of the patient. He makes a grave mistake if he deprives the patient of a stay in the sanatorium and indulges in home care upon diagnosis and thus withholds from the patient the specialized skills in the sanatorium during the early, trying, and treacherous months in which the disease is fluctuating or coming under control. In this way he is depriving his patient of the everlasting values that may accrue to him from learning how to rest and relax. Also he allows the patient to continue to expose his family and intimates for months while drugs are taking effect.

The family physician who enters such work with knowledge and enthusiasm, who singles out cases for discretionary care and refined diagnostic work-up, who is concerned with his patients who have been discharged from the sanatorium and whose plight will prompt him to take an interest in the latter's economic needs will automatically and axiomatically have an office that will be a case-finding unit, and his attitude will be one that will help usher in that day when the disease will have become relatively unimportant.

2. *Responsibility of the Public Health Physician.*—Formerly patients either applied to the physician for help or were never diagnosed. As educational efforts have focused attention increasingly upon the fact that the disease is often silent or unobtrusive, patients have correspondingly submitted themselves to the doctor or have applied in increasing numbers for x-ray examination, irrespective of symptoms. This development, along with the accompanying decrease in numbers of tuberculosis cases, has brought us to the point where case-finding methods have become increasingly popular, with 2½ million people in the United States being x-rayed last year. Useful as this method has come to be, however, its application yields a progressively smaller proportion of active cases as the incidence of tuberculosis diminishes.

The family physician and the case-finding unit continue to uncover new cases of tuberculosis and should carry on in full swing, but neither meets the total need. As the incidence declines, concentrated areas of infection can be spotted more readily, and detected cases stand out more clearly so that the detective in the person of the health physician assumes added significance, for

he pin-points the source—or sources—of the new cases. In fact, submitting to mass surveys is elective and its value appears to be declining. Frequently, if the health officer is alert, relatively larger numbers of active cases are subsequently uncovered among the smaller groups of the population refusing to participate in the surveys than among the much larger group who go along with it. The public health physician is in a position to do something active. He follows through, knowing that the most productive yield will be in certain racial groups, under certain conditions of poverty, ignorance, and bad hygiene, and in certain age groups. He uses the tuberculin test more effectively, for he is aware that one-half the deaths are in people older than 45 and that elderly persons, sometimes crotchety and contrary and uncooperative, pass off as unimportant the cough they may "have had for years." He knows also that a study of the contacts or tuberculin reactors leads to the source case much more often than could have been possible even twenty years ago, when infection was too prevalent for such accurate delineation. Thus, he wisely seeks to ferret out the sources in the population and is alert to those circumstances in which the disease develops with facility. From present trends it may be postulated that, as the disease becomes less common, this type of detective search for sources will be our most productive weapon against the spread of tuberculosis. The health officer can team with the general practitioner and sanatorium physician more effectively to combat this disease.

3. *Responsibility of the Sanatorium Physician.*—Among the people who carry responsibility for tuberculosis control, the sanatorium physician stands in a unique position. He it is who must see, evaluate, treat, and advise the patient, must come to know his medical and sociologic problems, and must deal daily or very frequently with him for many months through the changing moods that grow out of changes of his disease or from the disturbing information which the last visitor brought. He has an obligation to the patient to sit down with him, not just once but as often as necessary, to explain in laymen's language a few of the basic facts about tuberculosis—the role of the bacillus, what symptoms the disease may produce, what course it is likely to run if treated and if neglected, why various forms of treatment have been introduced and what these do to the disease and to the body, and which regimen seems best for him. Usually the patient should be told the therapeutic schedule that has been set up for him. His questions can and should be answered with sympathy and patience and his fears thus allayed.

In proportion as the sanatorium doctor can convince the patient of his real interest and friendship, the latter will develop confidence, will yield more readily to recommended therapy, and will have an understanding and cooperative point of view that may lead to a cure. It is the doctor's job to convince the patient of the importance of teamwork in the treatment and to emphasize that the patient's chief responsibility is to cooperate and rest.

The physician in the sanatorium has more to do than give air, administer sedatives, and prescribe laxatives. That he needs to keep his basic knowledge refreshed and up-to-date on the fast-changing picture in therapy goes without saying, but his responsibilities extend further. Within the last year a sanatorium

physician arose in a meeting of a society of nationwide membership and stated that chemotherapy represented about the sum and substance of the treatment of pulmonary tuberculosis. It was almost as if he reckoned his medical job to be that of the custodian who held the patient is reasonable flesh and nutrition until the surgeon came to do his bit, and then was keeper for a few added months until discharge. Is the treatment of a complex, protean disease such as tuberculosis as simple and devoid of judgment as this implies? Such problems as cough and its control, pulmonary drainage and cleansing, more logical bases of bed rest, nutrition of the patient, and the psychologic implications of his response to hospitalization and to his disease all constitute a challenge and an opportunity.

Irrespective of its soundness, the approaching trend toward a shorter stay in the sanatorium, followed by continued application of drugs at home, will be realized in some degree. It lays a new and particular burden upon the physician in the institution. For, actually, discharge of the patient with protracted chemotherapy implies a sharing of responsibility in a cooperative venture with the family doctor or health officer.

Thus, these physicians have an important role in the treatment of the individual patient and each must share his experience and "know-how" with the other. This calls for mutual consultation and, if possible, frequent contact. The man in the institution must often take the lead in this effort and often will need to outline a tentative routine with (not for, if you please) the man at home. The latter knows a great deal about his particular patient, who has his own particular intellectual and emotional background, yet frequently he knows but little about the average course of tuberculosis and the effect of many subtle factors upon it. The sanatorium physician will want to see the discharged patient on frequent occasions, often with the family physician if possible so that, together, they may carry the patient along the path to restored health and gainful occupation.

The sanatorium physician is in the peculiar situation which requires continuing contact with patients, nurses, laboratory personnel, and other physicians in the institution, so it is part of his job to contribute to good family relationships, good interaction between departments, and to act as a stimulant to the wholesome atmosphere of the place. These functions are important—sometimes in the long pull as important as the actual medical care of the patient.

And, too, the clinical problems which arise daily may, if the physician has imagination and ingenuity, become the topics of investigative or research work of high order and important result.

4. *Responsibility of the Surgeon.*—Proverbially, the physician and not the surgeon has cared for the tuberculous, but it is well known by all that the coming of effective chemotherapy for this disease has enabled the surgeon to take a vital part in treatment. This fact exaggerates a tendency, long among us, to separate sharply medical and surgical care. These, of course, are distinctly different; yet, at the same time, they are so intimately related to the recovery of the patient that they become one in objective and intent. The patient who needs operation needs medical care also, and he who is under a medical regimen may

require surgical intervention. So surgeon and physician together must determine the patient's requirements and, together, endeavor to supply them. In one case this may require a prolonged period of bed rest without drugs (if this is possible any more); in another, early operation may be indicated or, in a third, there may be exacerbation, failure to respond, or long-term, unchanging status. In any event we cannot separate the patient into a person needing medical care and one needing surgical care; rather he is one standing in need of intelligent care—be it medical, surgical, or both. Both medical man and surgeon are physicians and, in so far as each can learn and know the broad import of the special skills required of the other, they together will contribute to better treatment.

The responsibility of the doctor in tuberculosis control, then, is a broad and challenging one, embracing all the skills and knowledge at his command and inviting him to enter into a branch of medicine which in the last decade has come to offer many opportunities for service and for development.

OL.
1
955

REHABILITATION OF THE PATIENT WITH CHRONIC RHEUMATOID ARTHRITIS

A Two-Year Study

Edward W. Lowman, M.D.

New York, N.Y.

From the Department of Physical Medicine and Rehabilitation, New York University–Bellevue Medical Center and the Goldwater Memorial Hospital

VOL
1
195

(Received for publication Feb. 2, 1955.)

RHEUMATOID arthritis, unlike most other chronic diseases, is a disease of low mortality occurring in high frequency and causing marked incapacity. It is thus a disease with the virulence to disable and cripple progressively, but devoid of the mercy to kill. Neither its cause nor cure is known, and treatment is empirical or palliative. As a consequence of all these factors, the morbidity of rheumatoid arthritis assumes major significance in the over-all chronic disease and geriatrics problem.

EXTENT OF THE PROBLEM

The extent of rheumatic disease in the United States is known only by crude estimate. The survey by the United States Public Health Service[1] in 1938 estimated that nearly 7 million persons were afflicted with rheumatic disease. More recently, in 1952, Woolsey[2] surveyed 25,000 households in sixty-eight sample areas of forty-two states and estimated that in 1951, 10,104,000 persons beyond the age of 14 years considered themselves afflicted with rheumatism, which by special definition excluded rheumatic fever and rheumatic heart disease. Of this group, 6,414,900 had had such a diagnosis confirmed by their physician. Further, approximately 2½ million of these persons had had to make some change in the type or amount of work which they had been carrying on.

Rheumatism is so frequent and so disabling that it leads all other diseases in crippling and in economic loss. In the United States alone rheumatism accounts for a loss of more than 92 million man-days per year and an annual economic loss of one-quarter billion dollars.[3] Every year 147,000 persons are invalided by the disease.

This study has been supported by a grant from the National Institute of Arthritis and Metabolic Diseases, United States Public Health Service.

Cortisone and hydrocortisone have been generously supplied by Merck and Company, Inc. Geigy Pharmaceuticals has generously supplied Butazolidin (phenylbutazone).

Since these figures cover the entire group of rheumatic diseases, one may make certain approximate computations to estimate the incidence of rheumatoid arthritis. Inasmuch as 25 to 30 per cent of patients seen in a general clinic for arthritics have rheumatoid arthritis and 10 to 15 per cent of patients with rheumatoid arthritis progress to severe crippling and invalidism, the importance of the problem is apparent.

In the past, rehabilitation for the severely disabled rheumatoid arthritic has been reserved largely for the patient with a "burned-out" disease process; it is generally agreed that rehabilitation for such patients may be a successful undertaking because of the static nature of the disability. On the other hand, the question of rehabilitation for the patient with active rheumatoid arthritis has been moot, for not only is rehabilitation at times unable to keep pace with the progression of the disability but it has been uncertain whether intense rehabilitation procedures might not accelerate the disease process. The advent of adrenal and pituitary hormones for the better control of the inflammatory process of rheumatoid arthritis stimulated interest in assessing the value of a combined medical and rehabilitation program. A major objection to such a trial has been the frequent occurrence of undesirable hyperadrenal effects which develop in patients treated with these hormones for a long time.

OL.
1
955

In an effort to clarify the problem, we undertook an informal study of a pilot group of patients four years ago. These were all severely crippled rheumatoid arthritics with active disease who were hospitalized and treated with a combined cortisone and rehabilitation program at Bellevue Hospital in New York City. As a result of favorable impressions gained from experience with this pilot group, a full-scale research project was initiated in December, 1951, at the Goldwater Memorial Hospital to study the problem more throughly. The present report is based upon findings of the first two years with this latter group of patients. To carry out this study, a research unit was established consisting of twenty-five beds staffed by a rheumatologist, a physiatrist, a psychiatrist, a psychologist, social workers, a nurse-coordinator, and occupational and physical therapy personnel.

EVALUATION OF PATIENTS FOR REHABILITATION

Rehabilitation has been defined as the third phase of medicine, i.e., third to (1) diagnosis and (2) specific therapy. This expanded objective entails not only immediate medical care but also the return of the patient to his place in society functioning within the limits of his disability but to the hilt of his capability. In most cases this responsibility is assumed by the physician treating the patient for his specific illness; only in cases of severe disability are special rehabilitation facilities needed. The patients chronically disabled by rheumatoid arthritis fell into the latter group. Special treatment procedures were necessary for such individuals utilizing a team of specialists to combine medical management with physical therapy, occupational therapy, remedial exercise, psychotherapy, psychologic testing and vocational counseling, and assistance from social workers.

In the initial evaluation of a patient crippled with rheumatoid arthritis many kinds of information were needed before a coordinated treatment program could be formulated. These are shown in Table I.

TABLE I. REHABILITATION EVALUATION: THE TOTAL PATIENT MUST BE EVALUATED BEFORE A TREATMENT SCHEDULE MAY BE ESTABLISHED

Medical history Physical examination Laboratory examinations Specialist consultations	=	Diagnosis and physical prognosis
Muscle test Joint range of motion Activities of daily living Speech and hearing evaluation	=	Functional capacity and potential
Psychologic testing Social survey Vocational testing	=	Economic potential
	=	Rehabilitation potential

Medical Examination.—First of all, there had to be an accurate evaluation of the patient's arthritic process. This required a detailed history and physical examination, laboratory examinations, roentgenologic examinations, and, where necessary, the opinion of other consultants. An indication of the extent, severity, and "heat" of the arthritic process was derived from these data. It was of interest that of 285 patients referred to the project as severely crippled rheumatoid arthritics, twenty-four, or 9 per cent, did not have rheumatoid arthritis and 5 per cent had rheumatoid arthritis but their disabilities were from causes other than arthritis and of graver significance to them than their arthritis.

Functional Evaluation.—In addition to this medical appraisal of the patient's disease, it was necessary to have other information which might modify the prognosis and treatment goal. The muscle status of the patient had to be evaluated to identify weaknesses. Individual muscles were tested and graded on a special chart originally developed and used for patients with poliomyelitis. This evaluation of muscle power was pertinent in the rheumatoid patient because joints are of no value to a patient if the muscles, which are the "guy wires" about the joints, are inadequate to stabilize and mechanize them. Next, joint ranges of motion were measured and recorded not only to determine total range of motion but also to localize the arc of limitation. The latter was of considerably greater significance from a functional standpoint than the former. A knee, for example, may have a range of motion of greater than ninety degrees and still be contracted in flexion by a crucial fifteen degrees, thus precluding adequate weight-bearing.

The patient was then evaluated with actual functioning testing. While estimates of muscle power and joint ranges of motion were valuable sources of

information and suggestive of functional capacity, the inference as to function was not always valid, for pain had to be considered as an additional limiting factor. Patients were therefore additionally tested in performance of activities necessary to self-sufficient living. These many activities are collectively referred to as Activities of Daily Living (ADL) (Table II). This list of more than 100 activities ranges from the simplest bed and personal hygiene activities to the most demanding ambulation and elevation activities, including traveling via public transportation. The degree of functional impairment was thus directly measured and scored, and with the muscle power chart and range of motion measurements, correlation of functional with physical deficiencies would then be assessed. The sum total of these gave an indication not only of deficiencies but also of the physical potential of the patient.

Psychosocial Evaluation.—Finally, the psychologic, the social, and the vocational assets and deficits of the patient had to be appraised. What was the social and economic status? How had the patient psychologically adapted to the disease and what was his goal in rehabilitation? What intellectual and vocational assets were there upon which to formulate future plans? These factors were as important to the success of rehabilitation as any of the more obvious medical aspects.

The sum total of all these data (medical, functional, and socioeconomic) was then used collectively to define an individual rehabilitation goal and to initiate treatment toward this end. Goals as established initially often were changed repeatedly throughout the course of rehabilitation, as dictated by unanticipated changes in the patient's progress.

For the severe arthritic, treatment programs usually had to be full-day ones, individually prescribed. They included treatment sessions with various physical therapy modalities, occupational therapy, remedial exercise, functional training for deficiencies in ADL, psychologic and psychiatric assistance, vocational guidance, and, where necessary, job training.

METHOD OF STUDY

In the first two years of the study, fifty patients disabled by rheumatoid arthritis were accepted for treatment with a combination of drug and rehabilitation therapy.

In the first year of the study certain arbitrary criteria for admission of patients were rigidly adhered to. Patients had to be within the 21- to 55-year age group. The rheumatoid process had to be active and progressive. Patients obviously in need of corrective orthopedic procedures were excluded. There could be no medical contraindication for hormone therapy, and the patient had to have severe functional impairment. To quantitate functional incapacity more precisely than the Class I through IV system of the American Rheumatism Association, patients were tested and scored in the performance of Activities of Daily Living. These functional activities were categorized into ten major groups and to each group a score of 10 per cent was ascribed. Arbitrarily, then, it was decided that a patient must score at least a 40 per cent deficiency after

TABLE II

New York University—Bellevue Medical Center
Institute of Physical Medicine and Rehabilitation
400 East 34th Street
New York 16, N. Y.

PHYSICAL DEMANDS OF DAILY LIFE FROM BED TO JOB

DATE_____

NAME_____ WARD_____ AGE_____ SEX_____

ADDRESS_____ APPARATUS_____

CAUSE_____ DIAGNOSIS_____

DISABILITY_____ DATE ONSET_____ DATE REFERRED_____

METHOD OF RECORDING TEST

☐ 1. If at the time of the initial testing an activity cannot be performed *independently*, leave the block blank.

☐ 2. If the activity can be performed *independently* fill in the block with Blue pencil.

☐ 3. If the activity is not essential for the person's physical demands, draw diagonal lines in the block.

METHOD OF RECORDING PROGRESS

☐ 1. When the activity can be performed *independently*, fill in the block in Red and indicate the date of accomplishment.

I. NON-WALKING ACTIVITIES

			DATE
A. BED ACTIVITIES			
1. Moving from place to place in bed			
2. Roll to right and then to left side			
3. Sitting erect in bed			
4. Turn and lie on abdomen			
5. Procure objects from night table			
B. HYGIENE (TOILET ACTIVITIES)			
1. Combing or brushing hair			
2. Brushing teeth			
3. Shaving or putting on cosmetics			
4. Washing hands and face			
5. Washing extremities			
6 Manipulating bed-pan			
7. Applying Urinal/Special Pants			
8. Taking shower			
9. Tub Bath			
10. Ability to dry self after shower or bath			
11. Adjusting clothing for toilet needs			
C. EATING ACTIVITIES			
1. Cutting meat			
2. Buttering bread			
3. Eating with fork			
4. Eating with teaspoon, tablespoon			
5. Drinking from glass			
6. Drinking from cup			
7. Stirring coffee, tea, etc.			

This is the first sheet of a four-page list of activities considered essential to independent living.

VOL. 1 1955

stabilization of the disease process on maintenance doses of cortisone or other types of antirheumatic medication to be retained in the "severe disability" group. This group, consisting of twenty-four patients, is referred to as Group I.

In the second year of the study less rigid criteria were adopted to obtain a comparison group. These patients had to have active and progressive rheumatoid arthritis. Problems primarily of an orthopedic corrective nature were excluded. Functionally the patient had to be sufficiently disabled to be unable to carry on a gainful occupation, either a job in the case of the male or housework in the case of the female. Patients meeting these revised criteria are referred to as Group II.

In total, twenty-four patients meeting the initial criteria and twenty-six meeting the revised criteria were treated for periods of time varying from seven to twenty-one months. All were hospitalized on the rehabilitation wards of the Goldwater Memorial Hospital during this study.

After acceptance into this study patients initially were totally evaluated medically, functionally, and psychosocially along the lines outlined previously. Following evaluation, drug therapy was started; in most cases this was oral cortisone or hydrocortisone therapy. When maintenance dosages of the drug were established, physical rehabilitation programs were started and patients were continued on both drug and rehabilitation measures. As an index of improvement, scoring in proficiency in ADL was done before any treatment, following maintenance antirheumatic drug therapy, and repeatedly throughout subsequent weeks or months of rehabilitation. In this way it was possible to measure increased function induced by drug and by combined drug and rehabilitation treatment for comparison and objective evaluation of treatments. Following discharge from the hospital, patients have been followed at regular intervals as outpatients in a follow-up clinic.

RESULTS

In the initial two-year period of the study, 285 patients were referred to the project as "severely disabled rheumatoid arthritic patients" in need of a combined medical and rehabilitation program. For various reasons 235 of these referrals were rejected (Table III). It is significant that twenty-four of these, or 9 per cent, did not have rheumatoid arthritis but were suffering with other processes ranging from psychogenic rheumatism to metastatic carcinoma of the prostate. An additional fourteen patients, or 5 per cent, had rheumatoid arthritis but were more precariously disabled from other concurrent diseases. A total of fifty patients were accepted for treatment. Of these, twenty-four were accepted the first year and met the rigid criteria for severe disability as mentioned previously. In the second year twenty-six were accepted who were less severely disabled but still sufficiently disabled to be economically dependent. For various reasons (Table IV) six of each group were dropped from the study, leaving a total of eighteen in Group I and twenty in Group II. It is of interest to note that of these twelve, three were dropped because of frank psychosis, one because of borderline psychosis, and one because of hysteria; these clinical states

were verified after admission to the hospital and precluded treatment with adrenal hormones. Three patients of the severely disabled group died during their hospitalization, one of cardiac failure, one of glomerulonephritis, and the third of an overwhelming septicemia; in no case were adrenal hormones contributory. For only one of the three had a rehabilitation program been started, and this had no relation to the fatal course of the pre-existing disease process. The vital statistics of the two groups of patients are tabulated in Table IV; it should be noted that the duration of the arthritis in the more severely disabled group was roughly double that of Group II.

TABLE III. BREAKDOWN OF REASONS FOR REJECTION

a.	Not sufficiently disabled	49
b.	Not rheumatoid arthritis	24
c.	Over or under age	21
d.	Did not desire therapy on project	20
e.	Miscellaneous	19
f.	No care indicated (and/or arthritis burned out)	17
g.	Disability not due to rheumatoid arthritis but has rheumatoid arthritis	14
h.	Cortisone contraindicated	13
i.	Poor psychologic risk (including low I.Q.)	12
j.	Receiving care at clinic or by doctor who felt patient would not benefit from Project	10
k.	Lack of motivation	7
l.	Problem primarily orthopedic	6
m.	Unable to evaluate	6
n.	Needed care but not cortisone	6
o.	Needed only cortisone or other arthritis clinic care	6
p.	Nonresident	4
q.	Not financially eligible	1
	Total	235

Since the ability of a patient with an active rheumatoid process to participate actively in an intensive rehabilitation program depends to a large extent upon the physician's ability to convert the progressive process into a static or partially static one, most patients required maintenance antirheumatic medication. In twenty-five of the thirty-eight patients the active arthritis process was treated with cortisone or hydrocortisone, while eight patients were maintained on phenylbutazone. In only five of the thirty-eight patients could the disease be satisfactorily controlled by salicylates alone.

The results of treatment according to the American Rheumatism Association's method of classification are shown in Table V. It will be noted that major improvements have taken place in the functional classes with appreciable shifts from Classes III and IV to Classes I and II, indicating improvement in performance of activities essential to self-sufficient living. This is shown more succinctly in Table VI. Using the direct method of testing, the patient's ability to perform activities essential to daily living and his rating deficiency on a percentile basis are more precisely apparent. The average deficiency in Group I

VOL. 1 1955

was 60 per cent initially, 53 per cent after establishment of maintenance antirheumatic medication, and, finally, 29 per cent after rehabilitation. Similarly, in Group II, the initial disability average was 30 per cent; after maintenance medication this had decreased to 13 per cent, and the final average after rehabilitation was 4 per cent.

TABLE IV. ARTHRITIC PATIENTS ACCEPTED

	GROUP I (SEVERELY DISABLED)	GROUP II (LESS SEVERELY DISABLED)
Total number	24	26
Separated from project for following reason:		
Death	3	
Psychotic	1	2
Prepsychotic	1	
Hysteria	1	
Senility		1
Voluntary		3
Subtotal	18	20
Sex:		
Male	8	7
Female	10	13
Average age	46.2 years	40.2 years
Average duration of disease	13.2 years	7.2 years

TABLE V. AMERICAN RHEUMATISM ASSOCIATION CLASSIFICATION. RESULTS OF TREATMENT OF THIRTY-EIGHT PATIENTS WITH RHEUMATOID ARTHRITIS BY A COMBINED MEDICAL AND REHABILITATION PROGRAM. PERIOD OF OBSERVATION SEVEN TO TWENTY-ONE MONTHS

	INITIAL		DISCHARGE		GRADE IMPROVEMENT
	STAGE	CLASS	STAGE	CLASS	
Group I (18 patients)	0 = I 1 = II 6 = III 11 = IV	0 = I 0 = II 5 = III 13 = IV	0 = I 1 = II 6 = III 11 = IV	2 = I 10 = II 5 = III 1 = IV	0 = I 10 = II 8 = III 0 = IV
Group II (20 patients)	3 = I 4 = II 12 = III 1 = IV	0 = I 6 = II 10 = III 4 = IV	4 = I 3 = II 12 = III 1 = IV	9 = I 9 = II 2 = III 0 = IV	3 = I 11 = II 5 = III 1 = IV

The final disposition of patients at the termination of the initial two-year period (Table VII) left only four patients hospitalized as custodial care cases; these were all patients from the severely disabled group. As would be expected,

the results were better with the less severely disabled than with those of the severely disabled group. Over-all, however, twenty-two of the thirty-eight patients were totally self-sufficient upon discharge and twelve partially self-sufficient. Eight patients were placed in jobs outside the home; this does not include those housewives who returned to their home capable of resuming their jobs as housewives nor does it include patients engaged in home work.

TABLE VI. FUNCTIONAL DEFICIENCY*

	GROUP I	GROUP II
Number of patients	18	20
Before medication	−60%	−30%
After medication	−53% (15 patients)	−13% (18 patients)
After rehabilitation	−29%	− 4%
Duration of treatment	339 days	223 days

*Direct functional testing of patients with rheumatoid arthritis according to method shown in Table II.

TABLE VII. FINAL DISPOSITION OF PATIENTS

	TOTAL	% IMPROVEMENT IN FUNCTION
Group I	18	
Discharged totally self-sufficient	7	+39
Discharged partially self-sufficient	7	+26
Custodial hospital	4	+20
Jobs	1	
Group II	20	
Discharged totally self-sufficient	15	+21
Discharged partially self-sufficient	5	+29
Jobs	7	

CONCLUSIONS

A total of thirty-eight patients with chronic rheumatoid arthritis have been hospitalized in the past two years and treated with a combination of medical and rehabilitation measures. Eighteen of these patients were severely disabled cripples, while twenty were less severely crippled. The average age of the severely disabled group was 46 years, with a disease duration of thirteen years. Of the less severely disabled group, the average age was 40 years, with a disease duration of seven years.

Of the eighteen severely disabled, fourteen have been discharged from the hospital. Seven of these are totally self-sufficient and one has been placed in a

job. The other seven have been discharged partially self-sufficient with an average increase of 26 per cent in functional capacity. Four patients of the group remain custodial hospital cases, but among these there has been a 20 per cent increase in function.

Of the twenty less severely disabled, all have been discharged from the hospital. Fifteen are totally self-sufficient, while the remainder are partially so, with an average increase of 29 per cent in function. Seven of these twenty patients have been job-placed in full-time work.

The introduction of broad rehabilitation measures provides hope that the patient who is crippled with active rheumatoid arthritis may be salvaged from total invalidism.

The patients engaged in the study are now being followed for an additional three-year period to examine the validity of this inference.

OL.
1
955

REFERENCES

1. The National Health Survey, Bulletin 6, Sickness and Medical Care Series, U. S. Treasury Department, Public Health Service, 1938.
2. Woolsey, T. D.: Prevalence of Arthritis and Rheumatism in the U. S., Pub. Health Rep. **67**:505-512 (June) 1952.
3. National Research Council Final Report, Committee for Survey of Research on Rheumatoid Diseases, Washington, D. C., 1949.

Review

SOME APPLICATIONS OF BIOMETRY IN THE COLLECTION AND EVALUATION OF MEDICAL DATA

HAROLD F. DORN, PH.D.

BETHESDA, MD.

VOL
1
195

(Received for publication Dec. 10, 1954.)

ALMOST every phase of the practice of medicine necessitates at least the rudimentary application of statistical ideas. To the practicing physician, this statement, at first thought, may appear to be sheer fantasy. Yet the physician who mentally weighs and evaluates the various symptoms of an ailment in relation to the age, sex, or other characteristics of a patient and in relation to his previous experience with similar signs and symptoms in order to arrive at a diagnosis unconsciously is performing statistical operations. So also is the surgeon who, on the basis of his past experience, decides to perform an operation in preference to relying upon some other method of therapy or decides that an operation is too hazardous for a particular patient.

Nevertheless medicine, among the disciplines in which enumeration and measurement are commonly employed, has been the least influenced by the developments in statistical theory and techniques of the past two decades. There are signs that this situation is changing. In recent years there has been a growing tendency to insist that claims for the efficacy of specific methods of therapy be subjected to critical examination and trial, as is the custom in many other disciplines. An increasing number of examples of the application of the modern principles of experimental design can be found in medical journals.

But when we turn from the subject of clinical trials to that of investigations of the etiology of disease we find the basic concepts of biometry and epidemiology still largely unrecognized. A common practice is to publish an account of 100 (or some other round number) consecutive patients with a particular disease who have been admitted to a single hospital or have been treated by the author. The term "consecutive" apparently is intended to imply that the cases are unselected and hence may be considered to be representative of persons with the disease being discussed.

The age, sex, and race of the patients are described without any suggestion that these may reflect the distribution of such characteristics in the population

from which the cases are drawn as well as the selective effect of the disease itself. Similarly the existence of precursory diseases and symptoms or the coexistence of other specific ailments is mentioned without reference to the frequency with which these occur either in combination with other diseases or by themselves.

That this is so should not be surprising since it reflects the training of physicians in the interpretation of individual case histories. For every case of disease there is some chain of preceding events. It is this sequence that the individual case method attempts to discover. Much of the discussion in textbooks of medicine under the heading of epidemiology or etiology is based on evidence assembled from case histories which have been analyzed without reference to the population from which the cases arose.

The epidemiologist, fully aware of the importance of studying the characteristics of individual cases, concerns himself primarily with the investigation of disease as a group phenomenon of human or nonhuman populations. He views the development of disease as the result of the interaction of an etiological agent or agents, a host, and an environment. In contrast to case studies or clinical investigations which are concerned primarily with the antecedents of disease in specific individuals, its prognosis, treatment, and end results, epidemiologic research involves the study of disease in groups or aggregations of individuals.

The study of disease as a group phenomenon is dependent upon the ability to identify and count or measure relevant characteristics of a disease, of the population which it attacks, and of the environmental conditions under which it develops. To do this successfully requires use of the concepts and techniques of biometry. In the remainder of this paper I shall discuss some applications of biometry in the collection and interpretation of medical data.

DEFINITION OF ILLNESS

Statistics of ill-defined, unique, or dissimilar events not only are meaningless but also may be misleading. The first step in the application of biometric methods is an unequivocal definition of the events to be counted or measured. Since we customarily define health as the absence of disease, the usual event to be observed is a case of illness. But how shall illness be defined?

An operational definition might be, "An illness is what a doctor diagnoses." But physicians' diagnoses clearly not only fail to include all conditions of ill health that exist in a population but also are subject to error in the cases to which they apply.

The distinction between the living and the dead is easily made. The distinction between the well and the ill is not clear-cut and cannot be made precisely, with certain exceptions, since the transition from health to ill health frequently is almost imperceptible.

The state of health of a population may be represented by a continuous scale ranging from perfect health to the complete absence of health, that is, death. Between these two extremes all gradations of ill health exist. There is no agreement on criteria by which the various degrees of ill health may be identified and differentiated. The following diagram indicates some of the

degrees of ill health. It does not purport to be anything more than a rough representation of a few of the conditions that might be considered in the preparation of a scale.

Perfect health	Conditions predisposing to disease—e.g., obesity	Latent or incipient disease	Early apparent disease	Far-advanced disease	Death

Problems of diagnosis are particularly difficult for the ailments commonly called "chronic diseases."[1] In any disease, the person affected passes from a state of freedom from disease, to that of manifestation of disease, and finally either to death or to freedom from disease. It is a characteristic of the chronic diseases that the transition from absence of disease to presence of disease frequently is insidious and imperceptible. Moreover the state of ill health may persist for years with alternating periods of exacerbation and remission or with a steady progression through various stages of disease until death intervenes. It is especially important when studying the chronic diseases to recognize the wide range of manifestation of a particular disease.[2] A person with no clinical symptoms but with an amount of glucose in blood above some arbitrary level may be said to have diabetes, and a person in diabetic coma also may be said to have diabetes, but these two different stages of the disease are not of the same significance in most medical investigations.

Whenever empirical investigations of morbidity are undertaken, it is necest sary to decide at what point on the scale of health any departure from perfec-well-being will be called illness. The categories or degrees of ill health that are generally recognized are:

1. Congenital or acquired defects, impairments, or diseases which cause no current disability. For example, hemophilia, visual or auditory defects, and orthopedic impairments.
2. Latent or incipient diseases usually unrecognized by the person affected. As a rule, these conditions can be discovered only by special laboratory tests. Examples are the earliest detectable stages of tuberculosis, diabetes, or cancer. Opinions may differ as to the stage at which these conditions should be diagnosed as the clinical disease of which they are the precursor.
3. Manifest disease. This category includes persons who recognize they are not well. In addition, it includes persons with a clearly developed pathologic condition, such as cancer, which can be readily diagnosed with standard diagnostic tests but which is unrecognized by the person affected.

CRITERIA OF ILLNESS

Even though it logically may be maintained that a disease exists whenever certain pathologic changes have occurred, irrespective of whether these have been observed, for practical purposes we can count only the illnesses which have

been identified. The three principal criteria for establishing the existence of ill health are (1) the opinion of the individual concerned, (2) clinical examination by a physician, and (3) diagnostic tests.

It is sometimes thought that diagnoses are made by diagnostic tests. Diagnostic tests, in general, do not yield diagnoses nor do they necessarily objectively establish either the existence or the cause of illness. A diagnostic test may be qualitative, resulting in a picture, for example, an x-ray film, which must be interpreted by a reader. Or a diagnostic test may yield a reading on a quantitative scale. Criteria for relating specified scale values to corresponding degrees of ill health must be established before the results of such tests can be interpreted in terms of illness. The degree of validity of the calibration of a test and the competence of the interpreter of the result will materially affect the number of cases of illness discovered in a population.[3,4]

OL.
1
955

Since we can count only cases of illness which have been identified, it is only natural that concepts about illness are altered as more sensitive diagnostic tests are developed. The number of cases of tuberculosis that could be identified in a population markedly increased when clinical diagnosis based on physical signs and symptoms was supplemented with skin tests, sputum examination, fluoroscopy, and x-ray examination. Growing emphasis upon prevention and early detection tends to push the stage at which a pathologic process is called a specific disease increasingly closer to the beginning of the process. The diagnosis of carcinoma in situ, given prominence by the development of a cytologic diagnostic test, illustrates this point.

The determination of a diagnosis is much more difficult than the identification of the existence of ill health. The ability to make a diagnosis depends on the existence of standard criteria by which clinical findings and the results of diagnostic tests can be evaluated. The making of a diagnosis is based on medical judgment and hence is fallible.

The practice of therapeutic medicine is based on the assumption that it is possible to dichotomize the population into the sick and the well and that the diagnosis of an illness can be made so that appropriate therapy can be instituted. Epidemiologic investigations proceed on the assumption that diagnoses inherently are of varying degrees of reliability. The identification of a pathogenic organism in a culture, by itself, does not establish the existence of a clinically recognizable disease. The epidemiologist is as eager as the practicing physician to have diagnoses of the utmost validity. But since he is concerned with the natural history of disease he seeks to learn as much as possible, even from diagnoses based solely upon clinical examination.

Investigators, notably pathologists, coming to the investigation of the etiology and natural history of disease in human populations are prone to believe that reliable studies can be made only from cases with a diagnosis confirmed by some method common in pathology, in particular by autopsy. Aside from the fact that those who die as a result of a disease represent, in most instances, only a part of those who are ever subject to the disease and that autopsies ordinarily are performed on only a portion of those who die, the epidemiologist is interested in the investigation of disease in living populations. Hence he usually regards

autopsy statistics as a means of judging the validity of diagnoses of living cases of disease rather than the primary source of data for study. This point will be discussed in more detail later.

MEASUREMENT OF THE RISK OF ILLNESS

In order to study the occurrence of a disease in relation to the characteristics of the persons attacked or to the environmental conditions under which it develops, some measure of its frequency is necessary. Three measures of frequency are commonly employed: (1) number of cases, (2) relative frequencies, and (3) rates.

The simplest and most obvious measure of morbidity is the number of diagnosed cases. The age at diagnosis, sex, race, occupational and residence history, marital status, and other relevant characteristics can be determined and studied. Much of the discussion of the etiology and natural history of disease in medical literature is based on this method of measurement.

Reflection should reveal that very little beyond the fact that cases of disease exist can be learned with certainty concerning the epidemiology of disease by this method of measurement. The demographic and social characteristics of persons with a specific disease reflect not only the selective effect of the disease itself but also the characteristics of the population from which the cases arise. The age and sex distribution of cases of influenza among university students differs radically from the corresponding distribution for residents of a home for the aged. Yet by turning the pages, almost at random, of many medical journals, one can find articles in which the author comments on the difference in the average age, sex ratio, or some other characteristic of his series of cases when compared with some other published series and speculates about the reasons for this difference without taking into consideration that the simplest explanation might be that the corresponding populations from which the two series of cases were drawn had different characteristics. Another common misinterpretation in medical articles is to assume that the most frequent or modal age group of a series of persons with a disease is the age of greatest risk of developing that disease.

Much can be learned from a careful study of a series of cases concerning the pathogenesis and course of a disease. But one cannot progress far in epidemiology without a more reliable measure of the risk of illness than a mere enumeration of the number of cases. Without a standard of comparison, such as the population from which the cases arose, it is impossible to determine, from the number of cases alone, whether a disease is more common in one population group than in another.

Next to the number of cases of disease, the most widely used measure of the risk of illness in medical writing is a relative frequency. In general, relative frequency of a disease is the ratio, often expressed as a percentage, of the number of observed cases to the number of cases of all diseases. Thus the relative frequency of cancer may be expressed as the percentage that the number of diagnosed cases of cancer is of the number of diagnosed cases of all diseases admitted

to the hospital where the cases were observed or as a percentage of the total number of autopsies, if the study is based on autopsy statistics. Correspondingly, the number of cases of a particular form of a disease may be expressed as the ratio of the total number of cases of the disease as a whole. For example, the relative frequency of cancer of the lung may be computed as the ratio of the number of cases of lung cancer to the number of cases of all forms of cancer.

$$\text{Relative frequency of lung cancer} = \frac{\text{Number of cases of lung cancer}}{\text{Number of cases of all diseases}} \times 100$$

or

$$= \frac{\text{Number of cases of lung cancer}}{\text{Number of cases of all forms of cancer}} \times 100.$$

A relative frequency is a more direct measure of the risk of illness than is the absolute number of cases since it utilizes a standard of comparison. However, as its name implies, it is a relative and not an absolute or independent measure of the risk of illness. For males in the United States during 1950, 1.8 per cent of the total deaths and 13.9 per cent of the deaths from cancer were attributed to cancer of the lung. As can be seen from this example, the relative frequency of a disease depends upon the standard with which it is compared.

The relative frequency of a disease depends not only on the frequency of the disease in question but also upon the frequency of other diseases in the comparison group. The relative frequency of fatal lung cancer among men—expressed as a percentage of all cancer—in England and Wales is about 24 per cent. Without additional information one cannot determine whether the higher relative frequency of lung cancer in England and Wales, when compared with the United States, is due to the fact that lung cancer in reality is more frequent or to the fact that some other forms of cancer are less frequent there than in the United States.

Whenever one wishes to measure the importance of a disease *relative to* or *in comparison with* some other disease or group of diseases, a relative frequency is an appropriate measure to use. But if one wishes, as is more commonly the case in epidemiology, to measure the absolute risk of illness, the third measure listed, a rate, should be used.

A rate, when used to measure the risk of illness, has three essential elements, (a) a population exposed to the risk of some illness, (b) an interval of time in which the rate is expressed, and (c) the number of cases of illness developing in the population during a specified period of time. Rates usually are expressed as the number of cases of illness per 1,000 or per 100,000 persons exposed to the risk of becoming ill per year. A rate, of course, may be expressed per day, per month, or in terms of any other appropriate unit of time. The unit of time in which the rate is expressed should not be confused with the time interval of

observation of the population. A population may be observed for six months, two years, or any other period of time. The rate computed from the observed data may be expressed in terms of a different interval, if so desired, by suitable adjustment.

$$\text{Death rate from coronary heart disease} = \frac{\text{No. of deaths from coronary disease per year}}{\text{No. of persons exposed to risk of dying from coronary disease}} \times 100{,}000.$$

As used here, the purpose of a rate is to express the likelihood of developing a disease, of dying from a disease, etc., in a specified population during a specified period of time. Since observed populations differ in size, a rate is expressed per 1,000 or per 100,000 in order to adjust the observed data to a common size of population. The unit of time in which the rate is expressed must always be specified since the probability of becoming ill obviously depends upon the length of time the population is exposed to risk.

The difference in interpretation of the three measures of illness cited here may be seen from the data in Table I. The modal or most frequent age of persons who died from cancer in the United States during 1950 was between 65 and 74 years of age. The relative frequency of cancer—percentage of all deaths—was greatest for the age group 45-64 years. The highest mortality rate occurred at 85 or more years of age. The smallest number of deaths and relative frequency was during the first year of life, but the risk of dying was least for persons between 5 and 14 years of age.

TABLE I. MEASURES OF THE RISK OF DYING FROM CANCER, TOTAL POPULATION OF THE UNITED STATES, 1950

AGE	NUMBER OF DEATHS FROM CANCER	PER CENT CANCER DEATHS OF ALL DEATHS	NUMBER OF DEATHS FROM CANCER PER 100,000 POPULATION
0– 1	274	0.3	8.7
1– 4	1,521	8.4	11.7
5–14	1,635	11.2	6.7
15–24	1,917	6.8	8.6
25–34	4,763	11.2	20.0
35–44	13,440	17.5	62.7
45–54	30,372	**20.5**	175.1
55–64	52,232	**20.6**	392.9
65–74	**58,277**	17.0	692.5
75–84	37,802	12.4	1,153.3
85 and over	8,371	7.2	**1,451.0**

Source: National Office of Vital Statistics.

Even though the risk of dying from cancer increases from late adolescence until extreme old age, the relative frequency of cancer decreases after age 65 due to the fact that the mortality rate for cardiovascular diseases increases more

rapidly than the rate for cancer. The more rapid increase in the mortality rate for cardiovascular diseases results in an increase in the relative frequency of these diseases. If we compute the relative frequency of each disease in the standard of comparison, the sum must be 100 per cent. Hence, since the total is fixed, an increase in the relative frequency of one disease must be counterbalanced by a decrease in that for another.

Relative frequencies often are misinterpreted in medical literature. Almost uniformly they are interpreted as an absolute measure of the risk of dying or developing a disease. Much of the discussion of the epidemiology of non-communicable diseases which appears in medical books and journals is erroneous due to the misinterpretation of relative frequencies. An example may make this point more clear.

Assume that the annual incidence rates for cancer of the lung, breast, and uterus are the same in two populations but that the incidence rate for all forms of cancer combined is smaller in Population A. For purposes of illustration, assume also that there are 80,000 persons in Population A and 150,000 in Population B. Any other numbers could be used equally well since the relative size of the two populations is immaterial. Columns 2 and 5 of Table II show the expected number of new cases of cancer per year and Columns 3 and 6 show the relative frequencies of the different forms of cancer in the two populations.

Even though the incidence rates for cancer of the lung, breast, and uterus are the same in the two populations, each of these forms of cancer, relative to all forms of cancer, is more frequent in Population A. Data similar to these are commonly interpreted in medical literature as meaning that cancer of the lung, breast, and uterus are more frequent in Population A, implying by this that the risk of developing these forms of cancer is greater. The relative frequency of cancer of the breast in Population A may be higher than in Population B because (a) the incidence of breast cancer in fact is higher in Population A or (b) the incidence of some other form of cancer is higher in Population B. Which of these possibilities is true cannot be determined in the absence of morbidity rates.

TABLE II. HYPOTHETICAL ILLUSTRATION OF THE DIFFERENCE IN MEANING OF RELATIVE FREQUENCIES AND RATES

PRIMARY SITE	POPULATION A			POPULATION B		
	INCIDENCE RATE PER 100,000 POPULATION	COMPUTED NUMBER OF CASES	PER CENT OF TOTAL	INCIDENCE RATE PER 100,000 POPULATION	COMPUTED NUMBER OF CASES	PER CENT OF TOTAL
	1	2	3	4	5	6
Total	200	160	100.0	300	450	100.0
Lung	20	16	10.0	20	30	6.7
Breast	60	48	30.0	60	90	20.0
Uterus	50	40	25.0	50	75	16.7
All other	70	56	35.0	170	255	56.6

INCIDENCE AND PREVALENCE RATES

As was pointed out previously, an individual passes from a state of freedom from a disease to that of having the disease and finally to death or freedom from the disease again. A variety of rates have been developed to measure these changes in health.[5] Two morbidity rates are of primary importance in studying the epidemiology of disease; these are incidence and prevalence rates.[6]

Unlike death, which occurs at a point of time, illness exists during an interval of time. A statistical measure of the amount of illness in a population must take into account the time dimension of morbidity. The illnesses that exist in a population during a given time interval may be classified with respect to the time of development as follows:

1. Existing prior to the beginning of the interval, continuing throughout the interval, and still existing at the end of the interval.
2. Existing prior to the beginning of the interval and terminating during the interval.
3. Arising during the interval and still existing at the end of the interval.
4. Arising during the interval and terminating before the end of the interval.

The measure of primary interest in studies of the etiology and the natural history of disease is a measure of the risk or probability of a disease developing in a specified population during a given period of time. This is called the incidence rate and is defined as follows:

$$\text{Incidence rate of illness} = \frac{\text{Number of new cases of disease developing during a specified period of time}}{\text{Number of persons exposed to the risk of developing the disease}} \times 100{,}000.$$

The incidence rate is computed from Categories 3 and 4 of the classification of illnesses listed previously.

In addition to the incidence rate it often is useful to have a measure of the total number of cases of a disease in a population during a given period of time. This rate is called the prevalence rate and is defined as follows:

$$\text{Prevalence rate of illness} = \frac{\text{Number of cases of a disease existing during a specified period of time}}{\text{Average number of persons in the population}} \times 100{,}000.$$

The prevalence rate is computed from the sum of the four categories of illness listed previously. Prevalence is represented by the sum of the cases of illness existing at the start of an interval and the new cases developing during the interval.

Both the incidence and the prevalence rates may be computed for any specified time interval. In the study of communicable diseases it has been customary to interpret prevalence rates as a measure of the number of sick persons at any instant of time. In practice, this measure usually has been a daily prevalence rate. There is no reason to confine the use of prevalence to daily rates, and in the study of noncommunicable diseases it is often expressed as an annual rate.

Due to the difficulties in making a diagnosis and to delay in case finding, it rarely is possible to compute an incidence rate as defined here. Strictly speaking, we compute a substitute for the incidence rate, which more correctly should be called the discovery rate, since it is based upon the number of newly diagnosed cases of disease.[7] For example, newly diagnosed cases of cancer will include the earliest cases that can be diagnosed with existing diagnostic techniques as well as far-advanced cases with generalized metastasis that could have been diagnosed several months or even years previously. Even though every case of a disease developing in a population is diagnosed prior to death, the age curve of incidence will be biased since delay in case finding results in an average age at diagnosis that is older than the age at onset. In practice, nothing can be done about this in the absence of definite information concerning the duration of disease prior to diagnosis except to compute separate incidence rates for newly discovered cases classified by stage of disease at diagnosis.

An incidence rate may be defined in several different ways, although each definition is based on the concept of new cases or attacks developing during a specified time interval. A person may have more than one independent attack or episode of illness during a given period of observation. Furthermore, two or more diseases may exist simultaneously or sequentially during a single period of illness. Consequently, morbidity rates may refer either to persons, to periods of illness, or to diseases.

Incidence rates may be based on the first or initial attack of disease during an individual's lifetime or on a separate attack during the period of observation. However they are defined, new cases must be differentiated from recurrences or relapses. For some diseases, such as measles, the initial attack gives immunity to subsequent attacks; for other diseases, such as bronchitis or malaria, there may be a series of acute disabling episodes interspersed with periods of quiescence. Still other diseases, for example, acute upper respiratory infections, may attack the same person repeatedly with apparent complete recovery between attacks. A disease such as atherosclerosis may show a slow, unremitting progression until death is caused by a sudden acute episode of coronary occlusion or cerebral hemorrhage. Since incidence may be defined in several different ways, its meaning in any particular study should be clearly specified.

SOME PROBLEMS OF SAMPLING

The primary purpose of epidemiologic investigations is to discover generalizations that will hold true for a population larger than that on which the investigations were based. In this sense all studies are based on samples. The word "population" is used here with a different meaning than in the previous discussion

of rates. There we were concerned with populations or groups of persons on which specific observations could be made. In discussions of sampling, the word "population" is used in the sense of the source or supply from which persons are selected for study as, for example, the population or entire number of persons with diabetes admitted to a hospital, only a portion of whom are selected for a given study. "Universe" also is used as synonymous with population.

There is another more general meaning of population or universe in sampling theory, which is important in studies of the etiology and natural history of disease. We may investigate the relationship between a diet deficient in certain vitamins and health with the intention not of merely recording the nature of this relationship for the particular group of persons studied or for some specific larger population of which these are a sample but of drawing inferences concerning this relationship that will hold true for any corresponding group of persons wherever or whenever it may exist. In this instance, the observations are regarded as a sample from a hypothetical or imaginary population or universe composed not of persons but of all possible relationships between a diet deficient in vitamins and health. In most analytic studies of cause and effect relationships we are interested in drawing inferences that will apply to some hypothetical population of this nature.

If valid inferences are to be drawn from a sample, two conditions must be fulfilled: (a) the sample must be representative of the population to which inferences are to be drawn and (b) the data collected must be capable of answering the hypothesis to be tested or the questions to which answers are sought. In general, it cannot be said that one of these is more important than the other. In any practical problem, if a choice is necessary, one may decide to give greater weight to the desirability of obtaining more reliable data with some consequent restriction in the generality of the population concerning which inferences can be drawn. On the other hand, there are those who assert that the representativeness of the sample is paramount and that this should be preserved even if it involves some sacrifice in the reliability of the data collected. However, decisions of this nature should be determined by the conditions of the particular problem under investigation, and no logical basis for an a priori judgment exists.

Neither condition should be relaxed without careful consideration of alternatives since confidence in the results of a study is thereby weakened. Too often there is a temptation to relax both of these conditions because it is cheaper, more convenient, or quicker to take advantage of readily available data which are of doubtful validity and represent no population that can be readily described. The old adage, "Whatever is worth doing at all is worth doing well," can find no better application than in investigations of the etiology and natural history of disease.

How can we determine whether or not a sample is representative of some population? By "representative" is meant that the same conclusions, within the limits of sampling fluctuation, would be drawn from an analysis of the sample as from an analysis of the entire population. There is no way of ensuring that any given sample, drawn by any method, will be representative, in this sense,

of a specified population. Nevertheless some of the various methods which have been devised to select representative samples are more reliable than others.[8]

At first thought it might seem that the way to obtain a representative sample would be to choose the members of the sample deliberately to be representative of the parent population, with respect to some characteristic, or to pick a "typical" city, county, or similar area and study the people living there. But experience has demonstrated that this procedure may lead to biased results since certain members of the population almost always will be passed over. In addition, there is no rational basis for computing an estimate of the sampling error.[9]

At the opposite extreme is simple, unrestricted random sampling, a method which removes every element of purposive selection from the choice of the sample. It is not easy to define "random" in a rigorous manner. It does not mean hit-or-miss or haphazard. It implies that selection is determined solely by the laws of chance, that is, by the interaction of a multiplicity of small, independent causes, no one of which consistently exerts a predominant effect. The human mind is an unreliable instrument for the direct selection of a random sample. Consequently, statisticians usually employ some mechanical device or use a table of random numbers that has been especially prepared for this purpose.

Simple random methods are based on the principle that the selection of the members of the sample should have no obvious connection with the characteristics of the members selected and that each member of the population should have the same known probability of inclusion in the sample. Randomization has the great advantage of eliminating personal bias in the selection of the sample and of permitting the valid estimation of the sampling error.

Although simple, unrestricted random sampling has many useful applications and is of wide generality, there are situations for which it may be an inefficient method of selecting a sample. Experience has demonstrated that it is not always wise to remove all elements of judgment or purpose from the method of selection of a sample. It sometimes is undesirable to give every member of a population the same chance of being selected. At other times it may be essential for the purposes of the study that some members of the population be deliberately included.

These considerations have led to the development of methods of combined systematic and random sampling or of combined purposive and random sampling. In every method an element of randomization occurs at some stage. These, as well as other methods of choosing random samples, are called "probability sampling" and are characterized by the fact that each member of the population has a known, nonzero probability of selection, although this probability need not be the same for all members of the population. The description of these methods is too complicated to be discussed here. For further information a modern textbook on sampling should be consulted.

These methods have been devised for drawing samples from known, existent, enumerable populations. They have been tested repeatedly by experience and have been shown to be reliable and valid. When the population is known, a sample can be examined to determine whether it is representative with respect to relevant

characteristics such as age, sex, race, social class, or marital status. This always should be done no matter by what method the sample has been selected, since there is no firm assurance that any specific sample, even one selected by random methods, will be representative, as defined here, of the parent population.

One must be very cautious about rejecting a probability sample on the grounds that it appears to be unrepresentative. It is true that a large number of different probability samples usually can be drawn from a given population. However, if successive probability samples are drawn until one is obtained that appears to meet some preconceived notions, purposive selection, in effect, has been substituted for probability sampling. At the same time the results of any sampling procedure should not be blindly accepted. Judgment is required, but personal prejudice, bias, or predilection should not be confused with judgment.

In many medical investigations it often is difficult to define a population clearly and then to select a representative sample from it. This is particularly true in studies of the etiology of disease or of the efficacy of therapeutic methods. For example, during World War II the bread used in Great Britain was prepared from wheat flour from which not more than 15 per cent of the constituents of the wheat kernel had been removed in milling and to which calcium had been added. Throughout the war period there appeared to be a general improvement in the nutritional status of the population in spite of strict food rationing. At the end of the War some persons believed that the use of high extraction flour had contributed to the maintenance of the level of nutrition and recommended that the practice be continued indefinitely. Others took the position that the usual white flour in use prior to the war, if suitably enriched with vitamins, iron, and other minerals, would be equally if not more nutritious. It was suggested that these differing opinions should be tested by a feeding experiment with human beings.

Obviously the experiment must be performed on a sample population. How should the population from which the sample is drawn be defined and how should the sample be chosen? This is an illustration of the concept of a hypothetical population or universe mentioned previously.

Ideally, we would like to draw our sample of observations of the relationship between nutritional status and bread prepared from flour milled with varying rates of extraction and supplemented with varying amounts of vitamins and minerals from a population containing all possible interrelationships. But practically, we must make observations on some particular group of persons in some specific locality. How can we be certain that inferences drawn from the observed sample can be generalized to a larger universe?

There is no definite answer to this question. The purpose of the study will determine the scope of the inferences to be drawn. But the particular sample of persons selected for study may greatly restrict the generality of the inferences. In this type of situation the most that can be done is to select a sample of persons for study, bearing in mind the purpose of the investigation, and then by an analysis of this sample attempt to describe the population it may possibly represent.

VOL. 1 1955

The study just referred to was carried out in two German orphanages on children who were below the normal height and weight for their age due to an inadequate previous diet. During the period of the study, growth was equally rapid among all groups of children, irrespective of the type of flour and appreciably more rapid than for normal children of the same age. The study concluded by stating, "It remains an open question, however, whether the flours were equally nutritious, or whether differences existed between them which were obscured by the exceptionally rapid growth rate of undernourished children when receiving a diet adequate in calories. . . . The conclusions drawn must be restricted to the conditions under which the scientific evidence was obtained and cannot justifiably be applied to the needs of populations in very different states of nutrition and with widely varying dietary habits."[10]

This example illustrates how the choice of the sample of persons to be studied may greatly restrict the generality of the inferences that can be drawn from the results of the study in comparison with the original objectives. This was recognized by the investigators and clearly pointed out in the report of the findings of the study just referred to. It is not difficult to find published reports in which this point has been overlooked.

It is true that every group of persons may be thought of as a sample from some population. A study of the characteristics of the sample may permit a general definition of the population from which it could have come. But this deduction remains subject to uncertainty and usually cannot be demonstrated objectively. Consequently, the results of studies based on samples from undefined populations cannot be regarded with the same confidence as the results of studies based on samples chosen by some random method from a defined population. Before any study is undertaken, careful consideration should be given to the question of whether the objectives can be answered by observations based on the sample of persons selected. If they cannot, the study should be abandoned, its objectives redefined, or a different sample selected.

METHODS OF DETERMINING THE INCIDENCE OF DISEASE

One of the first steps in the epidemiologic study of a disease is to measure the frequency with which the disease occurs among persons with different characteristics, such as sex, age, race, marital status, socioeconomic class, and occupation, or who are living in different environments. In other words, the objective is to measure the risk of the disease developing in different groups of the population.[11,12]

As was pointed out, the only valid method of measuring the risk of development of a disease is by determining the incidence rate. A count of the number of diagnosed cases will reveal whether or not the disease has been recognized, but the number of diagnosed cases, by itself, cannot be used to measure incidence. The number of cases provides only the numerator for the incidence rate; a denominator is necessary before the rate can be computed.

The most widespread method of collecting morbidity data in the past has been by compulsory reporting of cases by physicians. Incidence rates have

been computed by dividing the reported number of cases by an estimate of the appropriate population, usually obtained from census reports.

Universal compulsory reporting has been confined almost entirely to infectious and contagious diseases. In recent years the desire for information concerning the incidence of chronic diseases has prompted some persons to recommend the compulsory reporting of these diseases. Except in special circumstances, the compulsory reporting of chronic diseases is not likely to be useful.

Universal reporting, even for the more serious communicable diseases, rarely has been satisfactory. Reporting has been seriously incomplete, even though the information requested has been reduced to the barest essentials. But in spite of the fact that incidence rates computed from reported cases have been subject to large errors, the compulsory notification of cases of communicable diseases has been useful in the control of epidemics.

VOL. 1 1955

Chronic diseases, however, do not occur in epidemics. The incidence rate changes relatively slowly, if at all. There is no need for annual morbidity data solely for the determination of incidence rates. Universal reporting, whenever it has been even partially successful, has been required for administrative, legal, or public health purposes. This is true for the reporting of births, deaths, marriages, divorces, communicable diseases, and occupational diseases. Until the reporting of cases of chronic disease can be justified for reasons in addition to the morbidity statistics that may result, it is not likely to be a useful source of incidence data.[13]

One justification for the reporting of chronic diseases is the provision of health or medical services. Many chronic diseases persist for months or even years, during which there may be need for medical care or for periodic examination. A case register can be a useful device for keeping a record of such persons and of the services given. If the register includes all cases arising in an area, the size and composition of whose population is known, it can be operated so that it also will yield valid morbidity data.

The number of diseases for which a case register is a practicable method of collecting data for the computation of incidence rates is not large. The objective of most case registers, quite properly, is to assist in the medical supervision and care of selected types of cases. This selection, in itself, makes the register unsuitable as a source of information to determine incidence.

Another drawback to the use of universal reporting as a source of morbidity data is the necessity of basing incidence rates upon population statistics obtainable from census enumerations. Ordinarily such publications give the size of the population classified by age and sex and sometimes color or race. The study of the incidence of a disease must be restricted to those characteristics which are identified both in morbidity reports and in available population statistics.

An effective method of overcoming these difficulties is to conduct a complete or sample survey of a specified population.[9] By this method the population is first defined. The desired information is then collected for both the well and the sick, usually by questioning a member of each family. The population

exposed to risk of illness is defined as part of the process of obtaining information about illness.

Illness can be related to a greater number of demographic characteristics and environmental factors than if the basic population data must be obtained from census statistics. This is the greatest advantage of the population survey since it allows the investigator to design studies to answer specific questions.

The population survey has several disadvantages. It is expensive if a large population is covered, as must be the case if data for specific diseases which are relatively rare are desired. But the greatest drawback is the degree of reliability and completeness of the medical information it is possible to obtain from family informants. The most that can be expected is a report of the diagnosis given the family by the attending physician together with the length of time the person affected was ill and information concerning any resulting disability. For many chronic diseases this is insufficient. When studying the incidence of cancer, for example, we may wish to know the primary site, the histologic type, the stage of disease at diagnosis, and the method of establishing the diagnosis. A family informant cannot be expected to have knowledge of such facts.

One solution to this problem has been used successfully in studies of morbidity from cancer.[14] A defined area for which population statistics are available from the census of population is selected for study. Information concerning every case of cancer diagnosed or treated during a specified period of time, one year for example, is collected from every physician, hospital, clinic, and medical facility in the area. Copies of death certificates mentioning cancer as a cause of death are obtained from the local registrar of vital statistics. Sources of medical care outside the area are checked for records of residents who may have left the area for diagnosis or treatment.

Information concerning the name and address of each patient is necessary in order to identify duplicate reports and nonresidents of the area. By means of the date of first diagnosis, cases can be separated into those first diagnosed during the study period and those first diagnosed prior to the study period, for the computation of incidence and prevalence rates.

This method has two important advantages over the household population survey. Information concerning all of the cases in a relatively large population can be collected by interviewing a small number of informants. The number of separate informants for a population of 1,000,000 persons ordinarily would be not more than 1,500 to 2,000. Second, specific detailed information concerning the disease and method of confirmation of diagnosis can be obtained, subject to the limitation that this information must be ordinarily available for every diagnosed case.

In common with other methods which depend upon the census of population for the denominator of the incidence rate, the number of characteristics of the population of the area to which the cases of disease can be related is restricted to those available from census publications. In metropolitan areas which have been subdivided into census tracts, ecologic studies of the incidence of disease in relation to housing, educational attainment, income, and socioeconomic status can be made by grouping the census tracts by some index of these factors.

The method is likely to be most useful when restricted to studies of relatively infrequent diseases, most cases of which will come to the attention of a physician and for which a diagnosis not made by a physician would be unacceptable. Higginson[15] has recently initiated a study of cancer among the Bantu in South Africa by combining this method of collecting information concerning diagnosed cases with a sample survey of the population of the area studied to determine the size and characteristics of the population at risk.

Useful information concerning the incidence of disease often can be obtained from the study of special groups, such as the employees of an industry or the members of a group medical care plan, provided information concerning the size and characteristics of the group is available. A recent study by Morris[16] of coronary disease among transportation company employees in London is an excellent illustration of this method.

Other methods can be devised for measuring the incidence of disease in particular situations. The essential point to recognize is that both a numerator and a denominator are required to compute an incidence rate. All too frequently, in the past, the discussion of the incidence of disease has been based solely on the number of diagnosed cases, since this information is more readily available than corresponding information concerning the population at risk. Once the necessity for information about both the numerator and denominator is generally recognized there should be fewer but better studies of the incidence of disease.

VOL. 1 195

OBSERVATION AND CONTROLLED EXPERIMENTATION

In its simplest form the purpose of the investigation of the incidence of a disease is to measure the frequency of the disease in different population groups or, in other words, to estimate the proportion of a group with specific characteristics and living in a specific environment that will develop the disease during a given time period. This is preliminary to the study of why observed differences in incidence occur, that is, to the study of the etiology of disease. In common parlance, the study of etiology is the study of the causes of disease.

Nearly everyone has an intuitive feeling of what is meant by cause and effect. Our everyday acts are based on the assumption that we can draw inferences concerning the future from past experience. The most general basis for belief in the cause and effect relationship of events is the observation that they are sequentially related in time. The first event is then thought to be the cause of the second. Until recent times this method of reasoning was the principal basis for the belief in the causal connection of events. Even today this method is so widely used that it would be virtually impossible to find a person who has not acted on beliefs established in this manner.

The principal error that may result from the observation of naturally occurring events is the inference that two events which are merely associated in time are somehow interrelated. The development of the experimental method provided a rational although not infallible procedure for detecting this error and made possible the analytic study of the relationship between two or more factors under controlled conditions.

The essence of the experimental method is (a) the formulation of an hypothesis, based on existing knowledge, of the possible relationship between two or more factors, (b) the definition of the population to which the results of the experiment will be generalized, (c) the testing of the hypothesis under known conditions with as many as practicable of the known causes of variation controlled or incorporated in the experiment so that their effects can be measured and with the effect of possible unknown causes of variation balanced by means of random allocation of experimental subjects to the test and control groups, and (d) the formulation of inferences based upon the results. Before such inferences can be accepted as a valid verification or disproof of the original hypothesis, they must be confirmed by repetition of the original experiment. Reproducibility is an essential part of the experimental method.

The controlled experiment, without doubt, is the most powerful method that has yet been developed for analyzing the nature of the relationship between two or more factors. It is the foundation upon which much of modern science is built. The rational appeal of controlled experimentation has led some to believe that only by this method can reliable knowledge of causal relationships be obtained. This attitude is exemplified by Sinclair's[17] statement that "the use of the experimental method has brilliant discoveries to its credit, whereas the method of observation has achieved little. The observer must await the occurrence of the natural succession of events he wishes to study, and he is very apt to be mislead by the fallacy of *post hoc ergo propter hoc* or by the existence of a correlation without causality."

If this statement were true, the outlook for increasing our knowledge of the etiology and pathogenesis of disease would be bleak indeed. The history of science bears testimony that this statement greatly underestimates the achievements of the method of observation.[18] There are the descriptive sciences, such as paleontology, astronomy, meteorology, geology, botany, and zoology. In addition, much of the knowledge in the biologic sciences, including medicine, has been derived from the observation of naturally occurring events.

The essential difference between the methods of controlled experimentation and observation of naturally occurring events lies not in the nature of the derived information but in the degree of confidence we may have in the inferences drawn from this information. The greater confidence that we attach to the results of controlled experiment is based on (1) the ability to specify the conditions under which observations will be made, (2) the ability to control or measure directly the effect of major causes of variation in the outcome of the experiment, and (3) the possibility of equalizing the effect of unrecognized or relatively less important factors by the random allocation of experimental subjects to the test and control groups.

PROSPECTIVE STUDIES

In observational studies, the method most closely approaching controlled experimentation is the prospective or forward study. This method starts with a defined population, which is kept under observation in order to determine the

frequency with which some event occurs. The characteristics of the population which may influence the occurrence of the anticipated event can be defined and described at the beginning of the period of observation. This ensures that the same information will be available for those who develop the condition as for those who do not. A known population at risk is available for the computation of rates to measure the probability of occurrence of the observed event and to determine its importance.

This method has been employed in a wide variety of studies. The effect of untreated syphilis upon longevity, the subsequent development of prematurely born babies, the relationship between smoking habits and mortality from various diseases, the dental caries experience of children living in communities where fluorides have been added to the water, and the mortality experience of persons with various diseases and impairments illustrate some of the uses to which this method of study has been put.[19-24]

A prospective study need not proceed forward in time from the moment it is first conceived. Studies are prospective whenever a population can be defined as of a specified time and the event to be studied can be observed after this time. Provided the necessary records were available, one could reconstruct a population as of some past date and study its experience with respect to the occurrence of some event after that date.[25] The necessary records for such studies often are not available.

RETROSPECTIVE STUDIES

A prospective or follow-up study starts with a population which is defined before some event occurs. A retrospective study starts with the event and then attempts to define the population from which the event arose.[26]

It is sometimes impossible or impracticable to conduct studies by the prospective method. More generally it is easier to make a retrospective study than it is to select a group of persons with a specific disease and attempt to identify the factors that influenced the development of the disease. It is easy to find a group of cases of a given disease; it is usually difficult to define the population from which the cases arose. Perhaps the most common occurrence in discussions of etiology in medical literature is, as Sartwell has so aptly said, a case series in search of a universe.

There are many reasons why this is so.[27] In the first place, the incidence of many diseases is so low that a very large population would have to be observed if a prospective study were carried out. A few years ago it was suggested that the babies born to women who developed rubella during the first trimester of pregnancy were likely to have certain congenital malformations. Even though more than 4,000,000 women become pregnant each year in the United States, the proportion of these who develop rubella during pregnancy is very small. This fact, coupled with the difficulty of locating the necessary thousands of women who have just become pregnant, has discouraged attempts to verify this suggestion by a prospective study. The few prospective studies attempted were inconclusive due to an insufficient number of pregnant women with rubella.[28]

Second, a long latent period may exist between exposure to a suspected causal agent and the appearance of an observable effect. The suggestion that smoking cigarettes may cause cancer of the lung is an example of this point. Information concerning the age at which most smokers begin to use tobacco regularly and the age incidence curve of lung cancer suggests that if there is a causal relationship, the effect may not manifest itself for as many as twenty to thirty years after smoking is begun. In addition, less than 5 per cent of males, smokers and nonsmokers combined, who reach 20 years of age would be expected to develop cancer of the lung during their future lifetime.

Faced with these and other difficulties, the usual procedure is to select a group of persons with the condition it is desired to study and by investigation of the past history attempt to reconstruct the cause of the condition. This is the method most widely used by physicians since the dawn of the practice of medicine to unravel the etiology of disease. Much of the discussion under the heading of etiology in medical articles is based on a description of the characteristics of a number of cases of disease. At best, this simple use of the retrospective method may lead to fruitful clinical impressions which may be tested by more carefully planned investigations. At worst, and the worst is all too frequent, it results in erroneous and misleading conclusions.

VOL.
1
1955

SOURCES OF DATA FOR RETROSPECTIVE STUDIES

The most commonly used source material in retrospective studies is individual case records. The investigator assembles the clinical records of patients seen during some time interval and tallies the frequency with which specified characteristics appear in the history. This probably is the most unproductive method yet devised for studying the natural history of disease.

The primary purpose of a case record is to document the basis for a diagnosis and the choice of treatment, to describe the treatment given, and to record the outcome. As a basic document for epidemiologic research it has three chief defects: (1) it may be incomplete, that is, the desired information is not available for each patient; (2) there may be lack of uniformity in recording data for different patients; and (3) more detailed information may be available for especially interesting cases, that is, those with unusual complications or combinations of conditions.

In taking a clinical history, a physician attempts to elicit facts thought to be significant in view of the symptoms presented by the patient. If he has a hunch about the probable diagnosis, his questions reflect his ideas concerning conditions likely to precede this disease. If he questions a man with cancer of the lip, he may inquire about pipe smoking or exposure to the wind and sun; a woman with cancer of the cervix will be questioned concerning her pregnancy history; a woman with cancer of the breast may be queried about lactation history but not about pipe smoking or exposure to the wind and sun. The tendency is to question each patient only about the conditions thought to be related to the existing symptoms or disease. Consequently, many clinical

histories merely reflect pre-existing ideas about the disease in question. This fact makes clinical records of limited value for studies of etiology.

Many discussions of the association of two or more conditions are based on data derived from autopsy statistics. The objective fact of the existence of the conditions usually is well substantiated. It is not the simultaneous existence of the two conditions that is in doubt; it is the interpretation of this association that is uncertain. This uncertainty is due to the selective character of autopsy statistics.

Rarely is it possible to autopsy every death. Equally rarely are the cases chosen for autopsy selected by some random method. Frequently an autopsy is performed because the case was unusual in some respect or some member of the staff is especially interested in all cases with a particular disease.

Even if the selection of cases to be autopsied were completely random, the interpretation of autopsy protocols still would be clouded with uncertainty. The occurrence of two disorders in the same person may make him more likely to seek hospital care and hence more likely to appear in autopsy statistics. Persons with one or more of the disorders studied may not be represented in the hospital population in the same proportion as in the total population. A final difficulty with autopsy statistics by themselves is that dead men tell no tales. The history of an autopsied case must be obtained from existing records.

Many, but not all, of the problems that arise in attempting to analyze hospital clinical records or autopsy statistics can be avoided by the planned interview of patients. The problem to be studied and the information required are decided upon before any cases are selected for study. Cases are selected according to some plan and every patient is asked the same questions by a trained interviewer. But even this method, which ensures that equivalent information will be available for every case, is inadequate, as a simple example will make clear.

Suppose that the history of 100 persons with diabetes reveals that ten also have a history of cholecystitis. Shall we conclude that cholecystitis is a precursor of diabetes? Clearly, without knowledge of the proportion of persons without diabetes who also have a history of cholecystitis, we have no way of deciding whether the proportion observed among diabetics is unusually high or low. In other words, a standard of comparison, usually called a control group, is required.

CHOICE OF A CONTROL GROUP

The primary function of the control group is to provide a basis for evaluating both the assumed explanation of the observed effect and other alternative explanations. The method of use of a control group for most retrospective studies differs in several important respects from the use of a control group in a planned experiment.

Basically, in a planned experiment, the subjects are divided into two groups by some random or combination of purposive and random process, one to be the test group, the other to be the control group. The purpose of random allocation

of subjects to the two groups is to ensure that each will be equally affected by extraneous or unrecognized factors that might influence the outcome of the experiment. The test group is subjected to some condition or treatment, and the effect is observed or measured in both the test and control groups.

In retrospective studies one starts with a group of persons each of whom shows the effect being studied, for example, a group of children with paralytic poliomyelitis, and then attempts to select a suitable control group in order to test whether some factor, for example, a recent history of tonsillectomy, increases the risk of being attacked by poliomyelitis. The control group is composed of children without poliomyelitis. It contains not only children who have not had a recent tonsillectomy but also those who have had a recent tonsillectomy.

The first step in selecting a control group is to define clearly the population of which it is to be representative. The control group should be as similar as possible to the test group in all relevant respects, except, of course, the presence of the disease to be studied. If this is not true, differences between the two groups may, in part at least, be due to the fact that the two groups do not come from the same population. This point should be kept in mind whenever the possibility of using data for the general population as a standard of comparison is being considered.

In the case of tonsillectomy and poliomyelitis, the population from which the control group should be a sample is made up of children who did not have poliomyelitis during the same period as the test group but who have been exposed to the same conditions relevant to the development of poliomyelitis, including tonsillectomy, as those who developed poliomyelitis. If the control group is selected entirely either from those who have never had a tonsillectomy or from individuals each of whom has had a tonsillectomy, it is valueless as a standard of comparison.

Two situations may arise when the control and test groups are compared. The supposed cause, for example, tonsillectomy, may appear only in the group with poliomyelitis. Whenever this occurs, one should investigate whether or not the two groups are, in fact, samples from the same population before accepting the relationship as valid, since a clear-cut difference of this kind is infrequent.

More generally, the supposed cause will appear in the histories of each group. The judgment of its importance is based on the percentage of children in each group with a history of tonsillectomy.

This method of judging the possible association between exposure to a supposed hazard and the subsequent development of some disease breaks down when both the control and the test group have had the same degree of exposure. For example, suppose that 100 mice have been given a weak carcinogen, which in the course of three months induces tumors in twenty. A person who knows nothing of the experiment examines the history of the twenty mice with cancer and the eighty mice without cancer. Finding that each mouse in both groups has been exposed to the same agent, he has no basis for assuming that this agent caused the cancer to develop. The fact that the suspected agent appears with

equal frequency in the history of both groups permits no conclusion concerning the existence or the lack of existence between it and the observed effect.

In most retrospective studies, the test group is the first N patients admitted to a hospital or group of hospitals after a given date. The characteristics of these persons with respect to factors, such as age, sex, race, marital status, and occupation, which might influence the development of the disease studied are not known until after the interviews are completed. Consequently, the population of which the control group should be a representative sample cannot be defined in advance. The best that usually can be done in this situation is to select the control group from the same source as the test group.

A few of the more important factors that influence the incidence of the disease may be controlled by matching the test and control cases with respect to these factors.[29] For example, the two groups can be made comparable with respect to sex, age, and race. If the number of persons interviewed is small and at least the total available test group or control population is large, this may be done by pairing. More generally, however, matching may more efficiently be carried out by stratification of the two groups, that is, by ensuring that the distribution by age and sex, for example, is the same for each group. Matching on more than a very few variables usually is impracticable due to the difficulty of locating a control subject. The measurement of the influence of other possible causes of variation can be done by cross-classification, with respect to these factors, of the observations made on the two groups.[30]

The choice of an appropriate control population should be carefully weighed before a study is started. It will be the standard by which the observations on the test group are evaluated. Unless it is well chosen, the conclusions from the study are likely to be invalid.

MEASUREMENT OF RISK IN RETROSPECTIVE STUDIES

In prospective studies a direct measure of the relationship between the supposed cause and the observed effect can be obtained by the computation of appropriate rates. This is not possible for retrospective studies. Instead the strength of the association must be measured by relative frequencies. In the most simple situation, the observations may be arranged as follows:

	POLIOMYELITIS	WITHOUT POLIOMYELITIS	TOTAL
With tonsillectomy	A	B	A + B
Without tonsillectomy	C	D	C + D
Total	A + C	B + D	N

The two proportions, A/(A+C) and B/(B+D), are compared. If A/(A+C) is greater than B/(B+D), the data are interpreted as indicating an association between tonsillectomy and poliomyelitis.

The first thing to observe about the table is that it proceeds from effect to cause and not from cause to effect. It shows the proportion of persons with tonsillectomy among those with and without poliomyelitis. What is desired is the proportion of persons with and without tonsillectomy who develop poliomyelitis. Why cannot the appropriate proportions, A/(A+B) and C/(C+D), be computed? Because (A+C) and (B+D) do not appear in the table in the same proportion as in the population from which the two samples were selected. If the sampling ratio were known, appropriate weights could be computed so that the desired proportions could be calculated, but this is rarely possible since the size of the population of which the control group is thought to be a sample is unknown.

If it can be assumed that the interviewed patients with poliomyelitis are a representative sample of all persons who developed poliomyelitis under the specified circumstances and that the interviewed patients without poliomyelitis are a representative sample of all persons who did not develop poliomyelitis, it is possible to compute a measure of the relative risk of developing poliomyelitis in the two groups of children with and without tonsillectomy.[32] Furthermore, if an estimate of the attack rate of poliomyelitis among all children is available, irrespective of whether or not tonsillectomy has been performed, an estimate of the absolute value of the attack rate in the two groups, with and without tonsillectomy, can be made.

If this assumption cannot be made, the data should be discarded. The assumption must be valid if any generalizations are to be drawn from the observed data: this holds true whenever relative frequencies or rates are computed.

Let $P_1 = A/(A+C)$ = the proportion of children with poliomyelitis who had a previous tonsillectomy, and

$P_2 = B/(B+D)$ = the proportion of children without poliomyelitis who had a previous tonsillectomy, and

X = the attack rate of poliomyelitis among children expressed per person.

Then the proportion of children with tonsillectomy who develop poliomyelitis is $P_1 X/[P_2 + X(P_1 - P_2)]$,

and the proportion of children without tonsillectomy who develop poliomyelitis is $(1-P_1)\, X / [(1-P_2) - X(P_1-P_2)]$.

If X is sufficiently small these formulas may be approximated by P_1X/P_2 and $(1-P_1)\, X/(1-P_2)$.

The ratio of the two attack rates can be estimated without a knowledge of X, the attack rate in the total population. When X is small, this ratio becomes $P_1(1-P_2)/P_2(1-P_1)$.

VALIDITY OF INFERENCES DRAWN FROM RETROSPECTIVE STUDIES

It was pointed out that retrospective studies start with an observed effect and attempt to discover the cause. In addition, the observations as made merely reflect the relationship of two events associated in time. In the example given, suppose that the children with poliomyelitis were mainly blondes while those without poliomyelitis were mainly brunettes. The trait, blue eyes and not

blue eyes, could be substituted for that of tonsillectomy and not tonsillectomy and a comparison of the proportions A/(A+C) and B/(B+D) undoubtedly would reveal that the former was considerably larger. Should it then be assumed that blue eyes predispose to or cause poliomyelitis? If not, how can we distinguish between situations when it is appropriate to draw causal inferences and those when it is not?

I know of no precautions that will enable one to distinguish infallibly between these two situations. Retrospective studies should be evaluated by an investigator who is thoroughly familiar with the subject, who is constantly aware of the possibility of mistaking concomitant association for cause and effect, and who is aware that some unrecognized factor may explain the observed relationship. Nevertheless, these are some precautions which may decrease the likelihood of accepting temporal association as evidence of a causal relationship.

First, however, it is well to keep in mind the types of relationships that may be observed. If A denotes a supposed cause and B the observed effect, these may be related in the following ways:

1. A is always followed by B
    a. B is always preceded by A
    b. B is sometimes preceded by A
2. A is sometimes followed by B
    a. B is always preceded by A
    b. B is sometimes preceded by A
3. A is never followed by B
    a. B is never preceded by A

The following questions should be considered before a retrospective study is undertaken:

1. What is the hypothesis to be tested?
2. What specific information is required to verify or disprove this hypothesis? Can this be obtained with sufficient reliability to warrant starting the study?
3. How would the study be conducted if it were possible to do it by controlled experimentation? What factors should be controlled?
4. How shall the sample of test cases be selected?
5. Can a suitable control group be found? How shall it be chosen?
6. What biases may exist as a result of the choice and method of selection of test and control groups? How widely can the inferences drawn from the results be generalized?
7. What other hypotheses might account for the observed effect?

Due to the inability to define and enumerate the population from which the test and control groups are chosen, the inability to select the two samples by some random method so that the effect of unrecognized factors will be equalized between the two groups, and the inability to control the circumstances under which the observed events take place, one should assiduously search for hidden

variables which might account for the observed relationship and check the original hypothesis in as many ways as possible.

Doll and Hill,[31] in their study of the use of tobacco and cancer of the lung, analyzed the results not only by amount of smoking, duration of smoking, method of smoking, type of tobacco, and type and site of the cancer, but also compared the test and control groups with respect to a variety of factors which might influence the observed association between smoking and lung cancer. The controls were selected to match the cancer cases with respect to sex, age, and place of interview. In addition, Doll and Hill compared the two groups with respect to occupation and social class, place of residence, exposure to different forms of heating, history of previous respiratory disease, residence near gas works, and the possible bias that may have resulted from the interviewer knowing the diagnosis of the persons interviewed.

If the relationship persists when the data are analyzed from widely different points of view, confidence is increased that the relationship is in fact as observed and does not result from the action of hidden variables.

After the data have been analyzed from as many points of view as possible, the following questions should be considered:

1. Are the inferences drawn from the study consistent with what is known about the disease being studied?
2. What questions remain to be answered before the observed relationship should be regarded as causal?

Finally, prior to publication, the manuscript should be given to one's severest critic.

These precautions, combined with a thorough knowledge of the subject, clear thinking, and strong skepticism, may help to avoid some of the more obvious pitfalls in retrospective studies. In the last analysis, it must be possible to reproduce the results, preferably under different circumstances, before they can be regarded as valid. Reproducibility does not necessarily establish validity, since the same mistake can be made repeatedly; but without reproducibility an observed relationship becomes merely an isolated historical event and adds nothing to accumulated scientific knowledge.

REFERENCES

1. Sartwell, P. E.: Problems of Identification of Cases of Chronic Disease, Milbank Mem. Fund Quart. **31**:234-238, 1953.
2. Sartwell, P. E., and Merrell, M.: Influence of the Dynamic Character of Chronic Disease on the Interpretation of Morbidity Rates, Am. J. Pub. Health **42**:579-584, 1952.
3. Garland, L. H.: On the Scientific Evaluation of Diagnostic Procedures, Radiology **52**:309-327, 1949.
4. Yerushalmy, J., Garland, L. H., Harkness, G. T., Hinshaw, H. C., Miller, E. R., Shipman, S. J., and Zwerling, H. B.: An Evaluation of the Role of Serial Chest Roentgenograms in Estimating the Progress of Disease in Patients With Pulmonary Tuberculosis, Am. Rev. Tuberc. **64**:225-248, 1951.
5. Stocks, P.: Sickness on the Population of England and Wales in 1944-47, Studies on Medical Population Subjects No. 2, 1947, General Register Office, England and Wales.
6. Dorn, H. F.: Methods of Measuring Incidence and Prevalence of Disease, Am. J. Pub. Health **41**:271-278, 1951.

7. Turner, T. B.: Studies on Syphilis in the Eastern Health District of Baltimore City. I. Principles Concerned in Measuring the Frequency of the Disease, Am. J. Hyg. **37**:259-272, 1943.
8. Cochran, W. G.: General Principles in the Selection of a Sample, Am. J. Pub. Health **41**:647-653, 1951.
9. Committee on Sampling Techniques in Public Health Statistics, On the Use of Sampling in the Field of Public Health, Am. J. Pub. Health **44**:719-740, 1954.
10. Widdowson, E. M., and McCance, R. A.: Studies on the Nutritive Value of Bread and the Effect of Variations in the Extraction Rate of Flour on the Growth of Undernourished Children, Medical Research Council Special Report Series No. 287, 1954 (Great Britain).
11. Gilliam, A. G.: Epidemiology in Noncommunicable Disease, Pub. Health Rep. **69**:907-913, 1954.
12. Expert Committee on Health Statistics, Third Report, World Health Organization Technical Report Series No. 53, 1952.
13. Schlesinger, E. R., Chase, H. C., and Le Boeuf, C.: Evaluation of Mandatory Reporting of Cerebral Palsy, Am. J. Pub. Health **44**:1124-1133, 1954.
14. Dorn, H. F., and Cutler, S. C.: Morbidity From Cancer in the United States, U. S. Public Health Service, Public Health Monograph. (In press.)
15. Higginson, J.: Study of Malignant Neoplastic Disease in Primitive Communities With Special Reference to South Africa, South African M. J. **27**:341-344, 1953.
16. Morris, J. N., Heady, J. A., Raffle, P. A. B., Roberts, C. G., and Parks, J. W.: Coronary Heart Disease and Physical Activity of Work, Lancet **265**:1053-1057, 1111-1120, 1953.
17. Sinclair, H. M.: Nutritional Surveys of Population Groups, New England J. Med. **245**:39-47, 1951.
18. Hill, A. B.: Observation and Experiment, New England J. Med. **248**:995-1001, 1953.
19. Rivers, E., Schuman, S. H., Simpson, L., and Olansky, S.: Twenty Years of Followup Experience in a Long-Range Medical Study, Pub. Health Rep. **68**:391-395, 1953.
20. Shafer, J. K., Usilton, L. J., and Gleeson, G. Untreated Syphilis in the Male Negro: A Prospective Study of the Effect on Life Expectancy, Pub. Health Rep. **69**:684-690, 1954.
21. Hammond, E. C., and Horn, D.: The Relationship Between Human Smoking Habits and Death Rates, J.A.M.A. **155**:1316-1328, 1954.
22. Doll, R., and Hill, A. B.: The Mortality of Doctors in Relation to Their Smoking Habits, Brit. M. J. **1**:1451-1455, 1954.
23. Society of Actuaries, Impairment Study, 1951. 1954.
24. Downes, J.: The Longitudinal Study of Families as a Method of Research, Milbank Mem. Fund Quart. **30**:101-118, 1952.
25. Brinton, H. P., Frasier, E. S., and Koven, A. L.: Morbidity and Mortality Experience Among Chromate Workers, Pub. Health Rep. **67**:835-847, 1952.
26. Dorn, H. F.: Philosophy of Inferences From Retrospective Studies, Am. J. Pub. Health **43**:677-683, 1953.
27. Dorn, H. F.: Some Problems in Sampling Human Populations, in Statistics and Mathematics in Biology edited by Oscar Kempthorne, Theodore A. Bancroft, John W. Gowen and Jay L. Lush, 1954, Iowa State College Press, pp. 241-249.
28. Hill, A. B., and Galloway, T. McL.: Maternal Rubella and Congenital Defects, Lancet **256**:299-301, 1949.
29. Cochran, W. G.: Matching in Analytical Studies, Am. J. Pub. Health **43**:684-691, 1953.
30. Greenberg, B. G.: The Use of Analysis of Covariance and Balancing in Analytical Surveys, Am. J. Pub. Health **43**:692-699, 1953.
31. Doll, R., and Hill, A. B.: A Study of the Aetiology of Carcinoma of the Lung, Brit. M. J. **2**:1271-1286, 1952.
32. Cornfield, J.: A Method of Estimating Comparative Rates from Clinical Data. Applications to Cancer of the Lung, Breast and Cervix, J. Nat. Cancer Inst. **11**:1269-1275, 1951.

VOL 1 195

OL.
1
955

Special Article

A BIBLIOGRAPHY OF INTERNAL MEDICINE: MALARIA*

ARTHUR L. BLOOMFIELD, M.D.

SAN FRANCISCO, CALIF.

From the Department of Medicine, Stanford University School of Medicine

(Received for publication Dec. 4, 1954.)

I'm awfully glad you love the beginnings of things—unless one does I don't believe one can get much pleasure out of the ends.

[Letters of Gertrude Bell, New York, 1927, Boni and Liveright, vol. 1, p. 81.]

THE rationale of this bibliography has been fully set forth in previous sections. The present installment, like the preceding one on tuberculosis, has presented great difficulties; the literature is overwhelmingly immense, many items are relatively inaccessible in Italian journals, and there are numerous offshoots into epidemiology, preventive medicine, protozoology, and pharmacology. As in previous chapters,† our guide has been the fact that this bibliography is compiled primarily for clinicians with general interests, such as medical students and internists. No apology is made therefore for slighting certain phases of the subject such as the public health aspects and some questions of immunity and of chemotherapy which the limits of space (and reason) make necessary. No doubt many significant papers have been entirely overlooked.

It should be recalled also that the two important decades—1880 to 1900—from the discovery of the parasite by Laveran to the culmination of Ross's work—was an era of bitter controversy, of disputed claims of priority, and even of accusations of frank plagiarism between the "Roman School" and first Laveran and later Ross. In Ross's memoirs (Ref. 39) there is even a chapter headed "Roman Brigandage." We have tried to deal fairly with this dispute, although it is easy to see how Laveran and Ross felt themselves ill-used.

The striking clinical features of the intermittent fevers have given them an ancient and honorable medical history which has been adequately dealt with

*Supported in part by a grant from the California Foundation for Medical Research.

†The preceding section is "A Bibliography of Internal Medicine: Tuberculosis," Stanford M. Bull. **12**:217, 1954.

by many writers. We have not attempted, therefore, to go back of modern times, that is to say, earlier than the nineteenth century.

For thorough discussions of every phase of the subject one should consult the monumental work on "malariology" edited by Mark F. Boyd (Philadelphia and London, 1949, W. B. Saunders Company). There are seventy chapters, each by an authority in the field, with comprehensive bibliographies. Good reference lists to the older literature are to be found in Laveran (Ref. 11), Marchiafava and Bignami (Ref. 18), Mannaberg (Ref. 20), Thayer and Hewetson (Ref. 22), and Ross (Refs. 36, 39).

MALARIA

Anemia—Ref. 5.
Avian malaria—Ref. 13.
Etiology—Refs. 6 through 9, 11, 12, 14 through 16.
Exoerythrocytic cycle—Refs. 46, 50.
Fertilization of parasite—Refs. 23 through 35.
Flagella—Refs. 19, 23, 24, 29.
General—Refs. 1, 2, 18, 22, 34, 39, 53.
Immunity—Refs. 32, 52.
Inoculation malaria—Refs. 10, 12, 38.
Melanemia—Refs. 1 through 5.
Modern chemotherapy—Refs. 40, 41, 44, 47, 51, 53.
Mosquito—Refs. 21, 25 through 27, 29 through 31, 33, 39.
Parasite—Refs. 7 through 9, 11, 12, 14, 16, 20, 48.
Pathology—Refs. 1, 2, 4, 48.
Prevention—Refs. 28, 32, 34, 36.
Quinine—Refs. 2, 43.
Staining of parasites—Ref. 17.
Transmission—Refs. 21, 25 through 27, 29, 31, 39.

VOL 1 195

1. Bailly, E. M.: *Traité anatomico-pathologique des fièvres intermittentes*, Paris, 1825, Gabon et Compagnie.

Reference is frequently made in the early literature to Bailly's book, mainly because of the pathologic studies, and he is often given credit for describing the characteristic malarial pigmentation of various organs. Actually the text is long and discursive, and the clinical pictures are too vague to have any useful meaning. There are numerous case reports with gross autopsy findings. These are not very satisfactory; in the "pernicious" intermittent fevers, arachnitis is a frequent diagnosis, often with "splenitis" or "gastroenteritis." The typical slaty or black pigmentation of the organs is mentioned only in passing; no special significance was attributed to it, and at best in an occasional case an organ was spoken of as dark (*foncé*) or grayish black (*gris-noirâtre*). In summary, the book serves best as a contrast for the advances in knowledge soon to come.

Bailly was familiar with the use of quinine sulfate but thought of it as "a specific sedative for the abdominal nervous system" (p. 426).

2. Maillot, F. G.: *Traité des fièvres, etc.*, Paris, 1836, J. -B.Baillière.

Maillot's book, eleven years after Bailly's, shows a good deal of advance. The disease observed in France, Corsica, and Africa is now recognizable as malaria; the clinical descriptions are excellent. However, there is as yet no precise differentiation of tertian and quartan or of other types. The deep pigmentation of organs is definitely noted; in one case the liver was chocolate colored, as was the spleen in seven cases (p. 285). In five instances of the comatose type the brain had a blackish tint (p. 287). However there was still no orderly synthesis of these observations and no realization that they had any special significance. Quinine was used along with other febrifuges; Maillot noted that where 5 grains of quinine sulfate was ineffective, 20 grains might prevent further paroxysms.

3. Charcot, J.-M.: "De la mélanémie, altération du sang par des granules et des corpuscles de pigment," *Gaz. hebd. méd. 4:* 659, 1857.

Pigmentation of certain viscera, especially spleen, liver, bone marrow, and brain, the mechanism of which is now clear, was long ago widely noted as an empirical fact in connection with intermittent fevers. We have already mentioned the observations of Bailly (Ref. 1) and of Maillot (Ref. 2). Richard Bright in his *Reports of Medical Cases* (London, 1831, Longman, Rees, Orne, Brown and Green, vol. 2, Case CI, p. 217, Pl. XVII) pictured a section of a brain which was "almost the color of black lead." Although this case is often quoted in the literature on malarial melanosis, there is no satisfactory evidence for such a diagnosis.

Meckel, a psychiatrist, is generally credited with being the first to note that brown pigmentation of organs in malaria was associated with an accumulation of pigment in the blood. His case report (Meckel, Heinrich: "Ueber schwartzes Pigment in der Milz und den Blute einer Geisteskranken," *Allg. Ztschr. Psychiat. 4:* 198, 1847) is, however, difficult to interpret. The patient had no clinical evidence of malaria and died in a psychotic state with weakness, edema, and ascites. The brain was chocolate colored, the capillaries were full of pigment granules, and the spleen was "very large" and dark brown. Meckel thought the pigment was formed in the spleen, whence it entered the blood. He performed various chemical tests but did not identify the pigment. Meckel later ("Körnigen Farbstoff in der Milz und im Blute bei Wechselfieberkranken," *Deutsche Klin. 2:* 551, 1850) pursued the whole matter further and brought more evidence that certain kinds of pigmentation of blood and viscera are characteristic of malaria. He now thought that the pigment came from blood which had been altered in the capillaries. Virchow went a step further ("Zur pathologischen Physiologie des Blut," *Virchow's Arch. path. Anat. 2:* 587, 1849) in reporting an autopsy on a case of intermittent fever in which he described pigment *in* blood cells which he thought were leukocytes. Whether or not any of these observers saw malarial parasites is hard to say; at any rate they had no inkling of their nature.

In Charcot's excellent article, modern in tone, is given an authoritative summary of current understanding of melanemia. The following facts were

accepted: Melanemia (literally, black blood) of marked degree occurred only in association with intermittent fevers. The pigment was brown or black, in small or large clumps, sometimes free in the blood or in organs, but more often in cells or "corpuscles" which were regarded as leukocytes, not as red cells. There was enlargement of the spleen of variable degree, less so of the liver. These and other organs often showed marked darkening; sometimes they looked black. The pigment was recognized as being related to the coloring matter of blood, and there was more when the number of red cells was greatly diminished. The association of acute comatose types of fever with plugging of the capillaries by innumerable "corpuscles" filled with pigment was clearly recognized, and the failure of quinine in some of these cases was thought to be explained by this. Hemorrhages from plugged capillaries were described. Exact relation of pigment to symptoms was not clear. Cachexia with dropsy and skin pigmentation was occasionally noted. Albumin and casts (Bright's disease) were observed in some cases. Pigmented corpuscles were seen in the blood of *living* people—often ten to fifteen to a "field of the microscope"; these were not thought to be red cells; but the decrease in red cells was recognized.

Where and how the pigment was formed were unknown; there was no idea of a parasite, and no understanding of phagocytosis by fixed tissue cells in organs. Charcot insisted that the pigmentation was not a primary disease; but he thought the study of these phenomena "should perhaps one day form an important chapter in the general history of malarial intoxication."

One wonders why none of the many workers on the subject suspected the presence of parasites. Poor microscopes, lack of staining methods, and study mainly of post-mortem material may be the explanation.

4. Frerichs, Fred. Theod.: *Klinik der Leberkrankheiten*, Braunschweig, 1858, F. Vieweg und Sohn, chap. VIII, "Veränderungen der Leber bei Intermittems," p. 325.

Frerichs, the great German physician of the mid-nineteenth century, gave a detailed and authoritative description, with beautiful plates, of the pathology of malaria. For Frerichs, all the lesions were probably the consequence of the melanemia. He described the pigment granules free, in cells supposed to be leukocytes, and in large cells probably of reticuloendothelial origin which he believed were lining cells of splenic sinuses. Frerichs thought that in malaria, as a result of febrile hyperemia, the normal tendency of red blood cells to be destroyed in the spleen was greatly exaggerated with massive pigment formation from excess of hemoglobin (*blutroth*) liberation. This in turn was carried to the liver and thence to other organs. Why melanemia did not occur in other fevers with splenomegaly was a puzzle. Frerichs' descriptions were authoritative since they were based on autopsies made by him during the study of an epidemic of intermittent fever in Silesia in 1854-1855.

5. Kelsch, A.: "Contribution à l'anatomie pathologique des palustres endemiques," *Arch. physiol. norm. et path. 11:* 690, 1875, 2 ser.

Although the occurrence of anemia in malaria had long been known, Kelsch seems first to have made systematic blood counts during life. He observed

VOL. 1 1955

initial rapid declines, often of as many as 1,000,000 cells a day, with less rapid later changes. He noted the extremely low counts in some cases of chronic paludism. The old discussions of melanemia are again gone through, but Kelsch did not get much further than Frerichs (Ref. 4). Again there was no suggestion of parasites. He thought that the pigment came from some obscure destruction of blood throughout the body, not only, or primarily, in the spleen. Even in the year when parasites were described (1880) the same absence of correct interpretation prevailed (Kelsch, "Contribution à l'histoire des maladies palustres. De la mélanémie," *Arch. gén. méd. 146:* 385, 1880).

6. Klebs, E., and Tommasi-Crudeli, C.: "Einige Sätze über die Ursachen der Wechselfieber und die Natur den Malaria," *Archiv. exper. Path. u. Pharmakol. 11*: 122, 1879.

The history of many infectious diseases, such as scarlet fever, influenza, poliomyelitis, and others, is that all sorts of bacteria were at first assigned a causal role. As one was shown not to be the true agent, another was blamed, and this error might be repeated many times before the true etiology was established. This was so with malaria; the early literature on bacterial claims is reviewed by Laveran (Ref. 11, 1891, p. 1 et seq.) and by Thayer (Ref. 22, p. 5). However, the findings of Klebs and Tommasi-Crudeli, supported by the authority of Klebs's name, seemed so plausible that they were generally accepted; nearly twenty years later, long after the discovery of the parasite, an editorial article appeared in one of the leading English medical journals stating the causal role of the bacillus of Klebs as a settled fact (Thayer, Ref. 22, p. 7).

Klebs and Tommasi-Crudeli, in brief, claimed that they could extract from the earth of malarial regions and from the air near the ground, bacteria which directly or in subculture, when injected into rabbits, produced a febrile illness with an intermittent type of fever resembling malaria, enlarged spleen, and at autopsy "abundant black pigment in the spleen as in people ill with malaria." They classed the organisms as bacilli and named them *Bacillus malariae.*

Klebs seemed to have all the necessary controls, and it is hard to understand how he could have erred so badly. A few months later (Klebs, E., and Tommasi-Crudeli, Corr.: "Studien über die Ursache des Wechselfiebers und über die Natur der Malaria," *Arch. exper. Path. u. Pharmakol. 11:* 311, 1879), the earlier brief statement was elaborated into a tremendous article of eighty-eight pages in which the details of the experiments leading to the same conclusions were given.* It is of interest to note the large number of workers, including

*In 1924 I recorded in my diary an anecdote told me by Professor Kraus during a visit to Berlin. "He (Kraus) began to speak of the malarial treatment of paresis and from this got on to malaria. It appears that before the discovery of the parasite, Klebs joined an Italian commission and did some field work, returning to Berlin with earth from the Roman Campagna. From this he grew a bacillus about which there was great excitement and which was regarded as the possible causal agent. Klebs inoculated two rabbits and left them overnight with two assistants (one of whom I suspect was Kraus) who were to watch and take their temperature. The assistants got a jug of beer and played Skat and finally went to sleep, and in the morning found the rabbits frozen to death. Klebs, then, in his lecture, much excited, said he couldn't demonstrate the temperature curves of the rabbits because they had died in the night *of malaria!*"

such men as Marchiafava, who thought they confirmed Klebs and Tommasi-Crudeli. Cuboni and Marchiafava, for example, began their paper on "New Studies on the Nature of Malaria" (*Arch. exper. Path. u. Pharmakol. 13:* 265, 1881): "Knowledge of the parasitic nature of malaria has received the most brilliant confirmation through the recent work of Klebs and Tommasi-Crudeli."

7. [Laveran] Presentation of a manuscript entitled "Note sur un nouveau parasite trouvé dans le sang de plusiers malades atteints de fièvre palustre," *Bull. Acad. méd. 19:* 1235, 1880.

It is difficult to say who first saw the malarial parasite. In systematic treatises various persons are mentioned whose descriptions of pigmented cells or bodies in the blood may have actually concerned the parasite, but primitive microscopes and inadequate technical methods without stains leave one quite uncertain. The bodies pictured by Virchow, for example (Ref. 3), bear no definite resemblance to the plasmodia.

VOL. 1 195

The actual discovery must be credited to Laveran. The popular notion, however, that as he gazed into his microscope the parasites were immediately evident and the whole matter clear is far from the actual record. Laveran's first communication was presented by proxy at the meeting of November 23 of the Academy of Medicine by M. Léon Colin, who stated that he had the honor to present for M. le docteur Laveran, stationed at the military hospital at Constantine (Algiers), a manuscript entitled "Note on a New Parasite Found in the Blood of Several Patients Sick With Malaria." The drawing which Laveran submitted is not reproduced, but three types of "elements" are described in the blood: "1. These are elongated, cylindrical, slightly curved like a crescent showing in their center a spot made up of blackish granulations. 2. Other round bodies, volume a little less than that of the red blood cells, sometimes immobile, sometimes moved by long tentacles which are attached to their borders. 3. Round elements much larger than the preceding doubtless because of swelling of the cells by imbibition of serum."

Colin went on to say in tactful terms that he doubted that these bodies were parasites, partly because of the noncontagiousness of malaria, and he proposed that a commission be set up to pass on the validity of similar future studies. One can see that Laveran was to have a rough road.

At the meeting of the Academy of December 28, M. Léon Colin again had a paper to present for M. Laveran entitled a "Second Note Relative to a New Parasite Found in the Blood of Patients Ill With Malaria" ("Deuxième note relative à un nouveau parasite trouvé dans le sang des malades atteints de la fièvre palustre," *ibid.*, p. 1346). The communication consists, however, mainly of a blast by M. Colin to the effect that Laveran's "parasites" were probably only leukocytes which had ingested pigment, a phenomenon common in the melanemia of malaria. Laveran, evidently on the defensive, now requested of the commission the aid of an expert histologist, M. Kiener of Val-de-Grâce, which in civil life was Laveran's own university.

8. Laveran, A.: "De la nature parasitaire des accidents de l'impaludisme," *Compt. rend. Acad. sc. 93:*627, 1881.

This time, at the meeting of Oct. 24, 1881, of the Academy of Sciences, Laveran spoke for himself. "There are present in the blood of malaria patients parasites which present themselves under the following aspects." He described four varieties: (1) Elongated crescentic bodies. (2) Spherical pigmented cells with filaments three or four times the diameter of a red cell in length. (3) Spherical pigmented bodies of irregular form which however are not pigmented leukocytes. (4) Smaller spherical, slightly pigmented bodies, some only one-sixth the diameter of a red cell. "The living nature of the spherical bodies inclosing mobile grains of pigment and armed with mobile peripheral filaments is unquestionable." Laveran also noted pigment granules and pigmented white cells in the blood. He stated that he had now studied nearly 200 patients with malaria. Most of those who showed no parasites had been treated with quinine. Parasites were found in no other disease. He noted that the parasites were most likely to be found shortly before the paroxysm of fever, but that in stubborn cases they were there all the time. He thought that between bouts of fever the parasites resorted to internal organs, especially liver and spleen. He noted in autopsies on malignant cases how the internal capillaries were packed with pigmented bodies. He concluded, "Attacks of malaria are produced by the introduction into the blood of parasites which appear in the various forms herewith described; it is because quinine kills these parasites that it cures malaria."

9. [Richard, E.], "Sur la parasite de la malaria," *Compt. rend. Acad. sc. 94:* 496, 1882.

Although Laveran's name now stands out, one must picture the tremendous ferment of activity which went on in the study of malaria in the 80's. For the most part doctors were skeptical or downright hostile to Laveran's claims. There were those who adhered to a bacterial cause like the bacillus of Klebs (Ref. 6); there were those who believed spores or molds were the etiological agent; and many still held to bad air or swamp emanations and did not believe in a *contagium vivum* of any sort. One can picture the patient hours spent at the microscope by faithful students as a "wild surmise" dawned that they might be looking at a living parasite. One of the earliest supporters of Laveran was his friend Richard, who worked in the hospital at Philippeville (Algeria), where there was much malaria. Richard read this paper before the Academy of Sciences on Feb. 20, 1882. The same remarks were published in the *Gazette médicale de Paris* in the issue for May 20, 1882 (Vol. 4, p. 252). Richard confirmed and strongly supported the discovery of the parasite which Laveran had named *Oscillaria malariae* because of the dancing appearance of the pigment granules. He examined fresh blood, unstained, and emphasized the fact, still true, that this is the best way to study the motility of the parasite. He differed with Laveran in one important point. Laveran thought the parasite was on the cell; Richard was sure it was *in* the red cell in which "it develops like a weevil in a lentil." He described the growth of the parasite at the cost of the host cell

until it is finally surrounded by a mere veil. Filaments are then thrust forth and finally the parasite bursts from its "shell."

10. Gerhardt, C.: "Uber Intermittens-Impfungen," *Ztschr. klin. Med. 7:* 372, 1884.

Inoculation malaria, later to become an everyday procedure in the treatment of paresis (Ref. 38), seems to have been first attempted by Gerhardt. A hypodermic syringe full of blood from the 37-year-old Brewer G. with intermittent fever was injected subcutaneously into the 39-year-old hysterical B. S. on Aug. 11, 1882. On August 22 there was chill and fever and splenic enlargement. The fever subsided after quinine. Gerhardt later repeated the procedure in another case. Nothing is said of parasites in this report, and there were no bacteriologic studies. After the parasite was recognized, attempts at transmission by inoculation were soon made; the early literature is well summarized by Thayer and Hewetson (Ref. 22, p. 35 et seq.), and very completely by Ross (Ref. 36, p. 66 et seq.).

11. Laveran, A.: *Traité des fièvres palustres avec la description des microbes du paludisme*, Paris, 1884, O. Doin.

In 1881, after the appearance of his first paper (Ref. 7), Laveran summarized the findings in a brief monograph entitled: *Nature parasitaire des accidents des l'impaludisme. Description d'un nouveau parasite trouvé dans le sang des malades atteints de fièvre palustre*, Paris, 1881, O. Doin. He continued, however, to work energetically on the subject, fighting a hard battle to convince others of the correctness of his views, which were later embodied in his book of 1884 dealing with every phase of malarial fever. In the preface he stated that it was in studying the mode of formation of the well-known pigmented elements in malarial blood that he was led to discover their parasitic nature. He continued somewhat nostalgically " . . . on the 6th of November 1880, . . . I demonstrated for the first time the presence of mobile filaments which adhered to the pigmented bodies and the living nature of which could not be in doubt. At this very moment I had the intuition that I was in the presence of the true parasites of malaria and everything I have done since has confirmed this impression." No indication of this sudden and dramatic revelation is given in his original reports (Refs. 7, 8). This book went through many editions in the next thirty years. Another important monograph appeared later (*Du paludisme et de son hématozoaire*, Paris, 1891, G. Masson), which was translated into English by J. W. Martin (*Paludism, by Dr. A. Laveran*, London, 1893, The New Sydenham Society). There is an excellent historical introduction. Although Richard (Ref. 9) believed that the parasites were intracellular, Laveran still felt that "our knowledge of the structure and composition of the blood-corpuscle in man makes it difficult for us to understand how these parasites can enter them. The probability is that they simply adhere to the corpuscles by pressing upon them. What makes this supposition more likely to be true is that the parasitic elements are often found in the blood in a free state" (p. 16). The latter statement is, of course, erroneous. Laveran recognized some difference between the

parasites in various types of malaria: decolorization of red cells in tertian, in contrast to quartan, diminution in size of parasitized red cells in quartan fever, and differences in segmentation in tertian and quartan fever. He was not certain that crescents were restricted to any special type of fever, as others had claimed. Although he quoted Golgi as believing that there are different races of malarial parasites, in the excellent plate at the end of the book all types—segmenters, flagellates, and crescents—are mixed together. Every phase of the disease including prophylaxis and therapy is discussed, and there is a section on avian malaria.

12. Marchiafava, Ettore, and Celli, Angelo: "Nuove ricerche sulla infezione malarica," *Arch. per le sc. med. 9:* 311, 1885. [A German translation of the manuscript was published simultaneously: "Neuere Untersuchungen über die Malariainfection," *Fortschr. d. Med. 3:* 39, 1885.]

OL.
1
955

Marchiafava's name has become so exalted in the history of malaria that it becomes important critically to analyze his work. Marchiafava stubbornly supported the etiological role of Klebs and Tommasi-Crudeli's "malaria bacillus," long after Laveran reported his work (Refs. 7, 8). Marchiafava and Celli in 1882 denied the parasitic nature of Laveran's bodies, although Laveran went to Rome and gave a demonstration of their presence in local cases* of malaria (Ref. 11, 1884, p. VIII, and 1891, p. 48). In 1883 Marchiafava and Celli ("Die Veränderung der rothen Blutscheiben bei Malariakranken," *Fortschr. d. Med. 1:* 573, 1883) hedged. Without any mention of Laveran they described what obviously were malarial parasites. "What now is the meaning of these bodies which we see (in malaria) within the red cells? Are they a regressive metamorphosis or are they foreign elements which have entered the cells? At the present time we lean to the latter interpretation and presume that we are concerned with a parasite . . . however none of this can be settled definitely as yet. We hope to clear the matter up by further work; meanwhile the first mentioned possibility that we may be dealing only with a regressive metamorphosis of red cells is still to be considered." This is evasive, to say the least. Laveran further said (Ref. 11, 1891, p. 48), "In a letter dated Rome, 9 April 1884, M. Marchiafava makes it clear that he and his collaborators have reached conclusions very different from mine. . . . 'We believe that the pigmented forms which you [Laveran] have described are nothing but degenerated and pigmented red cells. . . .' There can be no doubt on this point . . . four years after the publication of my work . . . they thought that the parasite of paludism was a micrococcus."

It would be interesting if one could reconstruct the medical politics of the time. Klebs and Tommasi-Crudeli were big names. The latter was head of the bacteriologic institute in Rome; Marchiafava worked in the same institution. There must have been great pressure to take sides in this controversy with a relatively insignificant doctor from Val-de-Grâce.

---

*Marchiafava and Celli described this visit and denied that Laveran demonstrated the parasites (Arch. ital. biol. *9*:308, 1888).

However, by 1885 Marchiafava and Celli seem to be suddenly committed to a parasitic cause for malaria. Marchiafava had become an expert on stained smears and, turning now to fresh preparations, he laid special emphasis on small ring forms which may be ameboid and which contain no pigment. He also described larger pigmented forms, for the discovery of which, as well as crescents, he gave credit to Laveran. In four among forty-two cases carefully studied, the flagella (*filaments mobiles*), also seen by Laveran, were noted. The activity of the flagella is vividly pictured. Marchiafava and Celli pointed out that filaments are not seen in stained blood. There is also a description of what seem to be segmenters and segmentation, although the drawing of *scissione* (division) is inaccurate and pictures a cell containing innumerable closely packed granules. Flagella and crescents were seen much less often than the other cells but were constantly present in pernicious fever. Marchiafava and Celli raised the question of whether segmenters give rise to a new generation of parasites and begged the question about the significance of filaments other than that they are living bodies. They felt certain that the small ameboid bodies are parasites which enter the red cells and destroy them.

They went on to describe attempts at transmission by injecting about 1 c.c. of blood freshly drawn from a malaria patient into a volunteer. The fact that they produced "typical" malarial fever, found the characteristic bodies in the blood, and made them and the fever go away with quinine—"all these circumstances make our hypothesis that these are living forms, probable."

Finally efforts were made to cultivate the parasite on ordinary bacteriologic media. In most cases the cultures remained sterile. "In others one obtained culture products which were morphologically identical with the initial forms [of the parasite]. . . . The multiplication of these bodies was especially evident in slide cultures of malarial blood."

It is obvious, then, that Marchiafava and Celli were still unclear about the significance of their findings, and in essence did little more than confirm Laveran's findings.

A little later Marchiafava and Celli ("Studi ulteriori sulla infezione malarica," *Arch. per le sc. med. 10:* 185, 1886; German translation from the manuscript: *Fortschr. d. Med. 3:* 787, 1885) elaborated their previous observations. They emphasized the small nonpigmented ameboid forms as their own discovery and followed them through to segmentation, which they thought was a method of multiplication to be followed by the invasion of the red cells by the liberated granules. They insisted that pigment was a product of the transformation of hemoglobin. It was in this paper that they suggested the name "Plasmodium" or "Haemoplasmodium," biologically unsound, but later universally adopted.* In this paper also the feud between Laveran and Marchiafava became overt. We cannot give the details of the controversy further, but Laveran's side is summarized in his book on *Paludism*, 1891, G. Masson, p. 48 et seq.

---

*For a full discussion of nomenclature see Covell, Sir Gordon, Russell, P. F., and Swellengrebel, N H.: "Malaria Terminology," Monograph Series, Geneva, 1953, World Health Organization.

13. Danilewsky, B.: "Zur Parasitologie des Blutes," *Biol. Centralbl. 5:* 529, 1885-1886.

Danilewsky, under the heading of "The Hematozoa of Birds," enumerated three forms. The description of the third form—clear rings which assume a spherical shape when more developed—could hardly have been anything else than "bird malaria." No analogy to the parasites of human malaria is suggested.

Many varieties of malaria in birds were later described; they have been of immense importance in the development of knowledge of the disease. The epoch-making observations of MacCallum on fertilization (Ref. 24) and of Ross on the cycle in the mosquito (Ref. 27) were made with bird malaria. Finally, during the war, innumerable potential antimalarials were screened by means of their effect on malaria of ducks.

OL.
1
955

14. Councilman, W. T., and Abbott, A. C.: "A Contribution to the Pathology of Malarial Fever," *Am. J. M. Sc. 89:* 416, 1885.

Although the prevalence of malaria in Italy and North Africa stimulated the early work of French and Italian observers, interest was soon aroused in this country. The foreign reports had hardly appeared before enthusiastic young American physicians began to study the subject in those regions of the Eastern seaboard where malaria was prevalent. However, most of this work was done in a spirit of confirming or disproving those foreign observations which were the subject of active controversy. Councilman and Abbott described in meticulous detail the lesions in two patients dead of "comatose malaria." They described the intense pigmentation of the organs and the small cerebral vessels choked with pigmented hyaline masses as if they were outlined in ink. The writers discussed the controversy as to a bacillary or protozoal cause and, like many early writers, begged the question. "We confess our inability to say what these hyaline bodies are."

In the same year Councilman ("Certain Elements Found in the Blood in Cases of Malarial Fever," *Tr. A. Am. Physicians 1:* 89, 1886) presented a critical discussion of the whole subject at the Association of American Physicians. "What the real nature of this body is—whether it is a parasite and the cause of malaria . . .time and further investigation will show." In the discussion Osler said, "I am not prepared to give a positive opinion as to the nature of these bodies. They look to me more like vacuoles or areas of hyaline transformation than definite organisms." One year later Councilman had described and accepted the parasites as the cause of malaria, and he clearly related different forms of the parasite to different clinical types of disease; "where the crescentic or elongated masses are found he has either some form of remittent [not intermittent] fever or malarial cachexia." ("Further Observations on the Blood in Cases of Malarial Fever," *Med. News 50:* 59, 1887.)

Another early American student who did important work was Sternberg, who reviewed the literature and described confirmatory observations of his own. He was inclined to accept the parasites as the cause of malaria (Sternberg, George M.: "The Malarial 'Germ' of Laveran," *M. Rec. 29:* 489, 517, 1886).

Finally there should be mentioned the later studies of Osler ("The Haematozoa of Malaria," *Brit. M. J. 1:* 557, 1887), who observed the parasites, now related them to malaria, and discussed their diagnostic value.

Many other Americans at about this time contributed useful studies on malaria, but they made no fundamental advances. All this work is reviewed by Laveran (Ref. 11) and by Thayer (Ref. 22).

15. Golgi, Camillo: "Sull' infezione malarica," *Arch. per le sc. med. 10:* 109, 1886. [Abstract in German by Carl Günther, *Fortschr. d. Med. 4:* 575, 1886.]

As everyone came to accept the parasitic cause of malaria, it was natural to inquire whether the different forms of fever were caused by special types of plasmodia. Laveran (Ref. 11, 1884, p. 196) said, " . . . one could raise the question of whether in the intermittent fevers the nature of the parasitic elements varies with the type of fever." He came, however, to no definite conclusion. Golgi took the first step toward distinguishing different types of parasites when he observed that in the quartan fevers prevalent in Pavia it took the whole length of the clinical cycle for the parasites which developed in orderly fashion to reach the segmented stage. He also observed that segmentation preceded a paroxysm and that immediately thereafter the parasites seemed to disappear from the blood only to return as the cycle was renewed. *He did not, however, in this paper positively claim that there was a special race of parasites in quartan fever*, although his pictures show beautifully that the segmenters always consisted of less than twelve merozoites. Golgi stated that the ameboid nonpigmented forms were probably distinct from the pigmented forms, although in a later paper (Ref. 16) he denied this.

A little later Golgi ("Ancora sulla infezione malarica," *Gazz. d. osp. 7:* 419, 1886; German abstract, *Fortschr. d. Med. 4:* 692, 1886) extended these meticulous observations to tertian fevers. The findings were analogous to those in quartan fever as far as an orderly development of the parasitic cycle is concerned; the differences between quartan and tertian parasites are well described. Golgi did not as yet positively assert that these were immutably different species.

16. Golgi, Camillo: "Sul ciclo evolutivo dei parasiti malarici nella febbre terzana," *Arch. per le sc. med. 13:* 173, 1889. [German translation: "Ueber den Entwickelungskreislauf der Malaria Parasiten bei der Ferbris tertiana," *Fortschr. d. Med. 7:* 81, 1889.]

By 1889 Golgi had clearly settled the fact that the various clinical forms of malaria were caused by distinct races of parasites. The well-known differences between the parasites of tertian and quartan fevers are systematically summarized. He noted that the early forms of tertian were more motile than quartan and that there was less early pigment formation; he observed the rapid decolorization of the red cells in tertian and the difference in number of segments in the two forms; he saw that the quartan merozoites were larger than the tertian and that the quartan parasites had larger pigment granules. He also described bodies

which may well have been sexual forms. In addition he said, "We now have reason to speak of still a third variety of malaria which is linked with the presence of the crescent-like bodies of Laveran." He noted the resistance to quinine of these bodies and pointed out the "irregular development" of the younger forms in this sort of [estivo-autumnal] malaria. The article is a classic and must be read in detail.

It is of interest that as late as 1891 Laveran denied the existence of several parasites and indeed of more than one form of malaria. "In 1884 I reached the conclusion that the different forms under which the haematozoa of malaria present themselves belong to one and the same polymorphous parasite; since then I have always defended this opinion" (Ref. 11, p. 126, 1891).

The cycle in estivo-autumnal fever was worked out by Marchiafava and Celli, whose preliminary report appeared in *Riforma med. 5:* 1281, 1889, and by Canalis, *ibid.*, p. 1443. The proof of a separate species was given by other Italian investigators (Gualdo, Tito, and Antolisei, Enrico: "Inoculazione delle forme semilunari di Laveran," *Riforma med. 5:* 1640, 1889), who gave a paretic person 2 c.c. of blood intravenously from a patient whose blood contained crescents. The recipient developed malaria with crescents.

An immense amount of further work on the details of the parasitic cycle in relation to clinical events was carried out in the next few years largely by the Italian school. These papers are reviewed in detail by Thayer and Hewetson (Ref. 22).

17. Romanowsky, D.: "Zur Frage der Parasitologie und Therapie der Malaria," *St. Petersb. med. Wchnschr. 15:* 297, 307, 1891.

No sooner had the parasites been generally accepted as living bodies than many workers began to study the details of their structure. The early work is reviewed by Thayer and Hewetson (Ref. 22, p. 45). Romanowsky in this paper described the important staining methods which still bear his name and reported important histologic studies of the malarial parasite.

An authoritative modern discussion is to be found in "Manual for the Microscopic Diagnosis of Malaria in Man" by Aimee Wilcox, National Institutes of Health (ed. 2), *Bulletin No. 180* (Revised), Federal Security Agency, Public Health Service, Washington, D. C., 1950. Both technical matters, such as the preparation of films, and the finer microscopic appearances of various forms of parasite are described.

18. Marchiafava, E., and Bignami, A.: *Sulla febbre malariche estivo-autumnali*, Roma, E. Loescher, 1892. [English translation by J. Harry Thompson, London, 1894, The New Sydenham Society.]

Following the original work of Marchiafava and Celli (Ref. 12), there was an immense resurgence of malarial study by the "young Roman" group of physicians. To be mentioned are Celli, Bignami, Bastianelli, Grassi, and Feletti and others. We cannot list the numerous papers, but their work dealt with the significance of the various parasitic bodies, the life cycle of the parasites,

their relation to the course of the disease, the effects of quinine upon them, and clinical classification and description of the various fevers in relation to the type of parasite involved.

Most of this work was summarized in this important book of Marchiafava and Bignami. They clearly described the different species of parasites and emphasized the spherical forms as the important part of the cycle, whereas Laveran had stressed the mobile filaments as the completed parasite. They perceived the relation between liberation of the "spores" by segmentation and the paroxysm, related the crescents to estivo-autumnal fever, but thought they were degenerated forms of the parasite, studied in detail the relation of quinine to inhibition of the parasite at various stages of the cycle, described accurately the malignant and comatose forms, and related them to plugging of the cerebral vessels by parasite-containing red cells. They had no idea of the sexual cycle and no inkling of the real meaning of crescents and flagella. There are excellent illustrations of parasites. This book is a storehouse of knowledge of the whole subject to date.

VOL. 1 195

19. Sakharoff, M. N.: "Recherches sur les hématozoaires des oiseaux," *Ann. Inst. Pasteur 7:* 801, 1893.

There had been great disputes as to the significance of the flagella of malarial parasites. Some, especially the Italian school, thought of them as degenerative dying forms; to others they seemed to be active bodies which must fulfill some important purpose. Sakharoff noted that the flagella of "crow malaria" contained nuclear material. "In our opinion the exit of chromatic filaments from the parasite and their energetic movements in their free state in the blood plasma is a fact which could be of general significance in cellular biology."

This was soon to be verified by the observations of Simond (Ref. 23) and of MacCallum (Ref. 24) and elaborated many years later by MacDougall (Ref. 49).

20. Mannaberg, Julius: *Die Malaria Parasiten*, Wien, 1893, A. Hölder. [English translation by R. W. Felkin, London, 1894, The New Sydenham Society.]

This scholarly book gives an authoritative summary to date of the knowledge of the malarial parasite based on Mannaberg's own work and on a comprehensive review of the literature. Mannaberg pointed out the widely differing views on crescents, and on the basis of morphology thought it likely that they are concerned with reproduction. From inoculation experiments in which the type of parasite in the injected blood was almost always found in the disease which was produced, he concluded that each parasite is a true species which does not undergo a transformation into other forms; there is a definite relation between the species of parasite and the type of fever. He had no idea how the disease is transmitted and was puzzled by the varying length of the incubation period. There are detailed descriptions of the parasites at all stages of the cycle and beautiful plates of both stained and fresh preparations.

21. Manson, Patrick: "On the Nature and Significance of the Crescentic and Flagellated Bodies in Malarial Blood," *Brit. M. J. 2:* 1306, 1894.

The story of the association of the mosquito with malaria is of great interest. One wonders over and over again how the causal relation could have been missed in view of the close association of mosquitoes with the circumstances of the disease.

King, in a much-quoted paper (King, A. F. A.: "Insects and Disease—Mosquitoes and Malaria," *Pop. Sc. Month. 33*:644, 1883), without any reference to the parasite, gave an orderly argument in favor of the mosquito being the transmitting agent. This is based on analogy to other mosquito-transmitted diseases—filariasis and perhaps yellow fever—and on the association of mosquitoes with malaria. Just how such transmission was brought about King, of course, did not know. Laveran (Ref. 11, 1891, p. 147) remarked that "perhaps these insects [mosquitoes] play a role in the propagation of malaria as they do in filariasis," but he pursued the idea no further. Similarly, others had mentioned more or less casually the possibility of insect transmission, but no one seemed inclined to perform any purposeful experiments until Manson. There is a good review of the early literature in Nuttall's monograph (Ref. 30, p. 75 et seq.).

Manson's paper seems to be the first in which any really impressive claim was made. Manson described and emphasized the forms later known to be gametocytes, described exflagellation, and reasoned clearly that the crescents and analogous spherical forms as well as flagella must have something to do with the propagation of the malarial process outside the human body. As there was no evidence that the parasites left the body in the excreta or through the skin, since the disease could be transmitted by injection of blood from malaria patients, since filariasis, a disease in many ways analogous to malaria, had been proved to be transmitted by the mosquito, Manson concluded that malaria also must be transmitted either by the mosquito or by some other suctorial insect. He postulated a cycle in the mosquito but thought it was not worth speculating further when experimental work could easily be done; this he hoped some one would achieve in the proper area, such as India. Manson's argument clearly included the flagella as an active and essential part of the cycle produced only outside the human host, and not as a degenerative stage of the parasite, as the Italian school held (Ref. 18). Manson did not know what happened in the mosquito or how the parasite returned to man; he did not state that it must be by mosquito bite, but at first rather thought it was by way of ingestion of water contaminated by infected mosquito larvae.

22. Thayer, William Sidney, and Hewetson, John: "The Malarial Fevers of Baltimore," *Johns Hopkins Hosp. Rep. 5:* 3, 1895.

This monograph is the first comprehensive and authoritative American work on malaria.* There is a thorough review of the literature with a bibliography

*The book by George M. Sternberg, "Malaria and Malarial Diseases," New York, 1884, W. Wood & Co., is definitely premodern in tone and there is no mention of the parasite which Laveran had described four years previously.

of 359 references dealing with all the early experimental work. The authors' own meticulous clinical and parasitological observations are next presented, and finally the classical drawings of malarial parasites by Max Broedel are here reproduced for the first time. Supplementary to this report is Thayer's *Lectures on the Malarial Fevers* (New York, 1897, D. Appleton & Co.). These works are definitive to date and still make exciting and valuable reading.

23. Simond, P.-L.: "L'évolution des sporozoaires du genre Coccidium," *Ann. Inst. Pasteur 11:* 545, 1897.

Simond, working with *Coccidium oviforme*, concluded that the flagellate bodies are male sexual forms. He did not actually observe conjugation but paved the way for MacCallum's work (Ref. 24). "The mobile stage of coccidium gives the most rational explanation of the flagella of malaria. . . . we should recognize the possibility of conjugation necessary for the production of a resistant form."

24. MacCallum, W. G.: "On the Flagellated Form of the Malarial Parasite," *Lancet 2:* 1240, 1897.

In this brief paper MacCallum, working with the malaria of crows (Halteridium), vividly described the formation of flagella and their conjugation with other sexual forms. "The granular forms lie quiet besides the nuclei and shadows of the red blood corpuscles which lately contained them but are soon seen to be approached by the flagella which having torn themselves away from the hyaline organism from whose protoplasm they were formed, struggle about among the corpuscles—one of them plunges its head into the sphere and finally wriggles its whole body into that organism." This epoch-making observation was confirmed in the case of a woman with estivo-autumnal fever, and finally resolved the numerous erroneous speculations of previous observers as to the significance of flagellated forms.

This paper was preceded by a brief note in the *Johns Hopkins Hospital Bulletin* describing conjugation ("On the Haematozoan Infection of Birds," *8:* 235, 1897), in which MacCallum gave Opie credit for describing the hyaline and granular sexual forms and for suggesting that the hyaline bodies gave rise to flagella (see Opie, E. L.: "On the Haemocytozoa of Birds," *J. Exper. Med. 3:* 79, 1898). In a later paper ("On the Haematozoan Infections of Birds," *J. Exper. Med. 3:* 117, 1898), MacCallum described the process of flagella formation and conjugation in much more detail, and there is a good drawing of the process. He had no idea, however, of the subsequent fate of the fertilized cell and made no mention of the mosquito. An analysis of previous views of the significance of flagellation is a useful part of this paper.*

*It is of interest that in historical accounts and in bibliographies (see, for example, Garrison, F. H., and Morton, L. T.: "Medical Bibliography," London, 1943, Grafton & Co., pp. 267-268), the name of E. L. Opie is often coupled with that of W. G. MacCallum in this discovery. Both were medical students working at the same time on avian malaria under the supervision of W. S. Thayer. Both presented their observations at the same meeting of the Johns Hopkins Hospital Medical Society on Nov. 16, 1896, and their preliminary findings were published in the Johns Hopkins Hospital Bulletin (8:51-54, 1897). Neither then nor later, however, did they publish jointly, and Opie to all intents seems to have

25. Ross, Ronald: "Observations on a Condition Necessary to the Transformation of the Malaria Crescent," *Brit. M. J. 1:* 251, 1897.

Ross threw himself into the controversy between Manson and Bignami (Bignami, Amico: "Le ipotesi sulla biologia dei parasiti malarici fuori dell'-uomo," *Policlinico 3:* 320, 1896) on the significance of crescents and of exflagellation. The Italian school, as we have said, looked on this as a degenerative involutional process; in other words, the flagellum was a dying form. Manson, on the other hand, thought the crescents and flagella represented a vital stage in an extracorporeal cycle of the parasite. Exflagellation presumably did not occur in human circulating blood but only after the blood had been altered in some fashion. Ross showed that if mosquitoes fed on a patient whose blood contained many crescents, there was a rapid change in the insect from crescent to sphere to exflagellation. In this paper he went no further in elucidating the cycle in the mosquito.

OL.
1
955

26. Ross, Ronald: "On Some Peculiar Pigmented Cells Found in Two Mosquitos Fed on Malarial Blood," *Brit. M. J. 2:* 1786, 1897.

In this paper Ross went a step further. In mosquitoes fed on patients whose blood contained many crescents he later found peculiar pigmented cells which suggested the presence of a parasitic stage. The specimens were sent to authorities in England; no one was certain, but Manson wrote, "Considering the peculiar grouping of the pigment in many instances, a grouping that forcibly recalls what one sees in the living malaria parasite, and the distinctness and regularity of the outlines of the bodies, I am inclined to think that Ross may have found the extracorporeal phase of malaria."†

27. [Ross] "The Role of the Mosquito in the Evolution of the Malarial Parasite. The Recent Researches of Surgeon-Major Ronald Ross, I.M.S.," *Lancet 2:* 488, 1898.

At the meeting of the British Medical Association, Manson reported the latest news of Ross's investigations on bird malaria (Proteosoma) in India. Ross had found that the pigmented bodies (Ref. 26) progress to "coccidia," which penetrate the stomach of the mosquito and protrude into the celom. "The coccidia now burst and . . . germinal vermicules which had been formed in

---

disclaimed the fertilization idea when he wrote, "After a while individual flagella not infrequently become detached and float away, continuing their active serpentine movements. After persisting for some time the flagella become less active and finally disappear, by exactly what process can not be said in every case" ("On the Haemocytozoa of Birds," *J. Exper. Med. 3*:95, 1898). Can it be that these brilliant medical students disagreed and that Opie did not accept, at least at the time, MacCallum's observation?

†In reading textbook accounts of these transactions (for example, Russell, P. F., West, L. S., and Manwell, R. O.: "Practical Malariology," Philadelphia, 1946, W. B. Saunders Company, p. 11), one gets the impression that Ross suddenly saw a parasite developing in the mosquito and that he "had doubt that he was looking at the malaria parasite." Nothing could be more misleading. Ross's ideas evolved gradually and painfully in the course of immense labor. "It would of course be absurd to attempt final conclusions as yet," "they seem to open the question of their being indeed the form of haemamoeba we are in search of," "the parasitic nature of the cells cannot finally be accepted until certain facts as to structure, sporulation and so on have been demonstrated" are phrases which Ross used in his paper. The whole story is vividly told in Ross's memoirs (Ref. 39, p. 225 et seq.).

its interior are set free in the body, blood and tissues of the mosquito. Then came . . . the discovery of those vermicules in the veneno-salivary glands of the mosquito." Finally Ross allowed mosquitoes to feed on infected birds and a few days later permitted these insects to bite birds whose blood was "void of any parasite infection." The birds became infected, and their blood was charged with parasites. "Thus the analogy between bird and human infection has only to be proved to establish that the mosquito is a carrier of malaria and an infector of man." The modest wording does not convey the full import of Ross's epochal clarification of the malarial cycle in the mosquito and the proof that the parasite was conveyed to man by bite and not by infected water.

This report was amplified a little later by a communication by Ross to the *British Medical Journal* ("Mosquitos and Malaria," *1:* 432, 1899); for Ross's other numerous communications on his discoveries the reader is referred to the bibliography given in his memoirs (Ref. 39).

While Ross is clearly responsible for tracing the cycle in the mosquito from gametocyte through to the salivary gland, he did not himself observe fertilization in the stomach blood of the mosquito, as Schaudinn later did (Ref. 35), and he did not have an opportunity personally to produce malaria in volunteers with human strains, as Manson did (Ref. 29), or to study the cycle of a human parasite in the mosquito until 1899 (Ref. 31).

28. Ross, Ronald: "The Possibility of Extirpating Malaria From Certain Locations by a New Method," *Brit. M. J. 2:* 1, 1899.

Laveran as early as 1891 (Ref. 11, p. 147) remarked, "Mosquitoes always abound in malarial regions and one notes soil drainage which suppressed the fevers, also suppresses the mosquitoes. . . . Perhaps these insects play a part in the propagation of malaria as they do in filariasis." He pursued the matter no further at the time.

But no sooner had the spread of malaria by the mosquito been proved than men of wide vision began to think about mass preventive measures. It is almost unbelievable how resistant to these ideas physicians in general were, as well as the bureaucrats without whose support nothing could be done. Ross, a pioneer in malaria control, told vividly about these difficulties in his memoirs (Ref. 39). In the present paper he discussed measures for eliminating the types of mosquito which carry malaria by draining and otherwise getting rid of the fresh-water pools in which they breed. This method of malaria prevention has, of course, been of immense importance—mosquito control—in the subsequent history of the subject. Koch, on the other hand (Ref. 32), was doubtful about the feasibility of eliminating mosquitoes, but on the basis of his work in Africa, proposed giving quinine to everyone infected with malaria, especially in the chronic cases with no fever, to prevent the mosquitoes from ingesting parasites. Manson (Ref. 29) laid emphasis on screening.

It is impossible in this bibliography to list all the important work on preventive and other public health measures in malaria. The brilliant work of Ross in Sierra Leone, the triumphs of Gorgas in Havana and in the Canal Zone, the useful efforts in the South Pacific in World War II require an extensive bibliography of their own. (But see Ref. 36.)

29. Manson, Patrick: "Experimental Proof of the Mosquito-Malaria Theory," *Brit. M. J. 2:* 949, 1900.

It is hard now to appreciate the general resistance in 1900 to belief in the mosquito theory of malaria transmission. Although A. Bignami ("The Inoculation Theory of Malarial Infection," *Lancet 2:* 1461, 1541, 1898) claimed in 1898 to have transmitted human malaria in Italy by mosquito bite, many people were unconvinced. In Manson's classical demonstration, infected mosquitoes were sent from Rome to London and there allowed on August 29 to bite a volunteer [Manson's son], who developed typical malaria on September 13 with tertian parasites in the blood and who was cured by quinine.* Conversely Drs. Sanborn and Low and Signor Terzi set up a mosquito-proof hut in an intensely malarial spot in Italy and lived in it from July to September without acquiring malaria "in marked contrast to their neighbors who were all of them either ill with fever or had suffered malarial attacks." Manson drew important conclusions from these experiments as to malaria control and prevention.

OL.
1
955

30. Nuttall, George H. F.: "On the Role of Insects, Arachnids and Myriapods, as Carriers in the Spread of Bacterial and Parasitic Diseases of Man and Animals. A Critical and Historical Study," *Johns Hopkins Hosp. Rep. 8:* 1, 1900.

This is a valuable review of parasites in relation to disease, including malaria, with useful bibliography of early work on mosquitoes and mosquito control.

31. Grassi, Battista: *Studi di uno Zoologa sulla malaria,* Rome, 1900, Tipographia della R. Academia dei Lincei. Also, ed. 2, 1901.

This classical monograph, handsomely printed, with beautiful color plates, deals with the work to date of the "Roman school" on every phase of malaria in relation to mosquitoes. The book has been bitterly attacked by Ross (Ref. 39) and others on the grounds of plagiarism. The Italians certainly took advantage of the fact that Ross had not had the opportunity to follow the cycle of *human* strains of malarial parasites in the mosquito to belittle his work.† However, we must let the individual reader adjudicate this dispute,‡ and at any rate the book is a storehouse of valuable information on mosquito transmission.

32. Koch, R.: "Zusammenfassende Darstellung der Ergebnisse der Malaria expedition," *Deutsche med. Wchnschr. 26:* 781, 801, 1900.

Koch, working in New Guinea, was one of the first to emphasize splenomegaly as an index of the degree of malarial infection in a locality. He found it

---

*This story is vividly retold in Ross's memoirs (Ref. 39, p. 417), and there is a complete summary of all the early mosquito inoculations in Ross (Ref. 36, p. 77 et seq.).

†A note in the British Medical Journal (*2*:608, 1899) entitled "The Cultivation of the Malarial Quartan Parasite in Anopheles" is a report of a telegram from Ross in Africa stating that he had studied the quartan cycle in the mosquito.

‡For a fair appraisal of the claims of Ross and of the Italian school see Nuttall, George H. F.: "On the Question of Priority With Regard to Certain Discoveries Upon the Etiology of Malarial Diseases," Quart. J. Micr. Sc. *44*:Part I, p. 429, 1900.

most common in children 3 to 6 years old who had had no evidence of active malaria for 2 to 3 years. By the time these children were 14 to 15 years old the splenic tumor was no longer palpable; other evidence of malaria had also disappeared. Koch thought all this due to an immunity reaction. Ross (Ref. 36, p. 127) discussed the matter systematically and introduced the term "*average spleen index.*" This has turned out to be very important from the epidemiologic and immunologic standpoints.

33. Giles, Geo. M.: *Handbook of the Gnats or Mosquitos*, London, 1900, J. Bale Sons and Danielsson. (Also ed. 2, rewritten and enlarged, 1902.)

During the whole of Ross's work on transmission the definition of the varieties of mosquito which were carriers of malaria presented one of the great difficulties. Mosquito classification in various areas was still inadequate, and Ross did not at first speak of Anopheles or Culex but described the insects as grey or brown, dapple-winged, and so forth. It was the gradual evolution of the work of many students which led to the conclusion that only certain strains of Anopheles were capable of transmitting malaria.

Giles's elaborate book was one of the early treatises in which mosquitoes were systematically described and those which transmitted malaria were identified. During the whole history of malaria studies it has been necessary to appraise the type of mosquito at fault in various regions where the disease was encountered. A key to the mosquitoes of the Australasian region was issued early in World War II (Knight, K. L., Bohart, R. M., and Bohart, G. E., Washington, 1944, *Nat. Res. Council, Division of Medical Sciences*, issued by the Office of Medical Information). A list of malaria-transmitting anophelines is given in Craig, C. F., and Faust, E. C.: *Clinical Parasitology*, ed. 4, Philadelphia, 1945, Lea & Febiger, p. 642; in Russell, P. F., West, L. S., and Manwell, R. O.: *Practical Malariology*, Philadelphia, 1946, W. B. Saunders Company, p. 587; and in Mackie, T. T., Hunter, G. W., and Watt, C. B.: *A Manual of Tropical Medicine*, ed. 2, Philadelphia, 1954, W. B. Saunders Company, p. 294.

34. Ross, Ronald: *Malarial Fever, Its Cause, Prevention and Treatment*, London, 1902, Longmans, Green & Co.

Although most of the principal facts about malaria were known by 1900, people in general were reluctant to accept the mosquito theory of transmission and its epidemiologic implications.

This little book by Ross is a classic. In simple terms all the fundamental knowledge about the disease ("containing full details for the use of travellers, sportsmen, soldiers, and residents in malarious places") is set forth, with special emphasis on the public health aspects, proper methods of using quinine, and so forth. "But there is no longer any reason at all why we should suppose that the infection of malaria rises from water or soil." "*It is not the germ which rises from stagnant water; it is the carrier of the germ which does so.*" [Author's italics.] The book did a great deal toward education of the public about malaria.

VOL 1 195

As early as 1900 Ross (*J. Soc. Arts 49:* 17, 1900) read a vigorous paper on "Malaria and Mosquitos." It concludes with a strenuous plea for the government to apply the obvious preventive measures for malaria. "In my opinion the period of probation has been completed and the time for action has arrived. Nothing will stimulate public action in this matter so much as will the authoritative support of our principal men of science."

35. Schaudinn, Fritz: "Studien über Krankheitserregende Protozoen. II. Plasmodium Vivax (Grassi und Feletti) der Erreger des Tertianfiebers beim Menschen," *Arb. a. d. k. Gsndhtsamte. 19:* 169, 1903.

In this thorough monograph on the life cycle of the tertian parasite, Schaudinn gave a minute and vivid description of fertilization in blood withdrawn from the stomach of mosquitoes who had fed on a patient with clinical malaria, thus amplifying the work of MacCallum (Ref. 24) and of Ross (Ref. 25). There are plates illustrating graphically the various stages of the process.

In this communication Schaudinn also described sporozoites (from mosquito bite) actively penetrating and invading the red cells. This work has never been confirmed, but belief in Schaudinn's observations held back for many years work on the exoerythrocytic cycle (Refs. 46, 50).

OL.
1
955

36. Ross, Ronald: *The Prevention of Malaria*, London, 1910, John Murray.

This is a comprehensive discussion of every phase of malaria prevention, definitive to date. There are contributions by a number of authorities in addition to Ross. A recent authoritative summary is that of Sir Gordon Covell ("Current Research Toward a Global Control of Malaria," *New England J. Med. 249:*125, 1953).

37. Bass, C. C., and Johns, Foster M.: "The Cultivation of Malaria Plasmodia (Plasmodium Vivax and Plasmodium Falciparum) in Vitro," *J. Exper. Med. 16:* 567, 1912.

From the time of the discovery of the parasite, attempts at cultivation were made. These early efforts, in spite of some claims of success, were all clearly failures. A number of interesting experiments were, however, made on survival of the parasites in leeches which had fed on malarial patients. O. Rosenbach ("Die Conservierung lebender Malariaparasiten," *Berl. klin. Wchnschr. 28:* 839, 1891) believed that such parasites remained alive at least forty-eight hours. N. Sacharoff ("Ueber den Einfluss der Kälte auf die Lebensfähigkeit der Malariaparasiten," *Centralbl. Bakteriol. 15:* 158, 1894) found parasites still motile in leeches which had fed seven days before. Thayer and his associates confirmed these findings (Ref. 22, *Lectures*, p. 27).

Bass in 1911 made a preliminary note ("A New Conception of Immunity," *J. A. M. A. 57:* 1534, 1911) on his cultivation work, but this is the definitive report. The technique is given in full. "In successive cultures the asexual parasites grow, segment, and form rosettes, which burst and give rise to merozoites, many of which enter new red blood cells in exactly the same manner as

they do in the body of man." It was important that *Plasmodium vivax* and *Plasmodium falciparum* cultivated in identical fashion remained distinct. The subject was carried further and its difficulties and complexities were discussed by Trager ("Studies on Conditions Affecting the Survival in Vitro of a Malarial Parasite [Plasmodium lophurae]," *J. Exper. Med. 74:* 441, 1941, and "Studies on the Extracellular Cultivation of an Intracellular Parasite [Avian Malaria]," *ibid. 92:* 349, 1950). Geiman and his associates (Geiman, Q. M., Anfinsen, C. B., McKee, R. W., Ormsbee, R. A., and Ball, E. G.: "Studies on Malarial Parasites. VII. Methods and Techniques for Cultivation, "*J. Exper. Med. 84:* 583, 1946) also have a valuable paper on the problem of cultivation. Hawking reviewed attempts to grow malarial plasmodia in tissue culture (Hawking, Frank: "Tissue Culture of Plasmodia," *Brit. M. Bull. 8:* 16, 1951). The upshot of all this work is that as yet there is no simple, reliable method of cultivating human malarial parasites through successive generations.

VOL. 1 1955

38. Wagner-Jauregg, J.: "Die Behandlung der progressiven Paralyse und Tabes," *Wien. med. Wchnschr. 71:* 1106, 1210, 1921.

The early inoculation experiments (Refs. 10, 12) were carried out purely to determine the transmissibility of malaria. Wagner-Jauregg, however, introduced inoculation malaria as a means of causing therapeutic fever in general paresis. While the subsequent huge literature (authoritatively reviewed by Moore, J. E., in *The Modern Treatment of Syphilis*, ed. 2, Springfield, 1941, Charles C Thomas, Publisher, p. 374 et seq.) deals largely with the results of treatment, valuable data have also been accumulated on the natural history of inoculation malaria (for example, Yorke, Warrington, and McFie, J. W. S.: "Observations on Malaria Made During Treatment of General Paresis," *Tr. Roy. Soc. Trop. Med. & Hyg. 18:* 13, 1924), and its value in the study of immunologic questions and as a means of testing therapeutic agents has been widely explored. Kopeloff, for example, noted that after seven months of transfer through patients a strain began to lose its capacity to produce gametocytes and became asexual (Kopeloff, N.: "Inoculation Malaria: Sexual and Asexual Strains," *Am. J. M. Sc. 197:* 800, 1930). A condensation of Wagner-Jauregg's article soon appeared in English ("The Treatment of General Paresis by Inoculation of Malaria," *J. Nerv. & Ment. Dis. 55:* 369, 1922).

39. Ross, Ronald: *Memoirs, With a Full Account of the Great Malaria Problem and Its Solution*, London, 1923, John Murray.

In the course of these fascinating memoirs Ross tells the intimate story of his malaria research. During the years in which this work was done (circa 1895 to 1900) Ross was a bitter, disappointed man. He had to work under the most adverse conditions, he was over and over frustrated by an incredibly stupid bureaucracy, and, when the discovery was finally made, presumably reputable scientists tried to steal his glory and a skeptical public was inert to the tremendous sanitary implications. Ross was essentially an idealist and a humanitarian. His discovery thrilled him not as a biologist but because it indicated the way

to elimination of a dread disease. As he pointed out, he was a sanitarian rather than a zoologist. The book is a mine of source material and should be read by everyone interested in malaria or public health.

40. Sinton, J. A., and Bird, W.: "Studies on Malaria With Special Reference to Treatment: Plasmoquine in the Treatment of Malaria," *Indian J. M. Research 16:* 159, 1928.

Plasmochin (pamaquine), a synthetic quinoline derivative, had been described as an antimalarial drug in 1926 by Schulemann and others (Schulemann, Weiner, and Mennini, Guglielmo: "Plasmochin, ein synthetisches gegen die malariainfection wirksames chinolinderivat," *Arch. Schiffs-u. Tropen Hyg. Beih. 3, 30*: 59, 1926). At that time the importance of differentiating action on the erythrocytic and exoerythrocytic stages of the parasite was not well understood, and as the suppressive and preventive actions of the drug were not outstandingly better than quinine and there were important toxic symptoms, its use was not pushed. Sinton and Bird first clearly demonstrated the action of Plasmochin in preventing relapses, that is, in producing radical cure. Sinton, Smith, and Pottinger continued these studies ("Studies in Malaria With Special Reference to Treatment. XII. Further Researches Into the Treatment of Chronic Benign Tertian Malaria With Plasmoquine and Quinine," *Indian J. M. Research 17:* 793, 1930) and showed that in situations where the relapse rate was 70 per cent or more in patients treated with quinine only the addition of Plasmochin reduced the relapses to 20 per cent or less. Other reports to the same general effect were those of James, Nicol, and Shute ("The Prevention of Malaria With Plasmochine," *Lancet 2*: 341, 1931), of Most, Kane, Lavietes, London, Schweder, and Hayman ("Combined Quinine-Plasmochin Treatment of Vivax Malaria: Effect on Relapse Rate," *Am. J. M. Sc. 212:* 550, 1946) working with Pacific strains of vivax, and of Berliner, Earle, Taggart, Welch, Zubrod, Knowlton, Atchley, and Shannon ("Studies on the Chemotherapy of Human Malarias, VII. The Antimalarial Activity of Pamaquine," *J. Clin. Invest. 27:* Part 2: 108, 1948) with experimentally induced vivax and falciparum malaria. Here again pamaquine had a relatively feeble suppressive effect, even in maximal doses. If the schizogenous cycle was interrupted by quinine, then the curative (antirelapse) effect was evident in vivax malaria.

41. Mauss, H., and Mietsch, F.: "Atebrin, ein neues Heilmittel gegen Malaria," *Klin. Wchnschr. 12:* 1276, 1933.

The writers discussed the composition and properties of Atabrine (quinacrine U.S.P.), which was soon recognized as a valuable antimalarial. Early trials (Peter, F. M.: "Ueber die Wirkung des Atebrin gegen natürliche Malariainfection," *Deutsche med. Wchnschr. 58:* 533, 1932; Napier, L. E., and Das Gupta, B. M.: "Atebrin: Synthetic Drug for Treatment of Malaria," *Indian M. Gaz. 67:* 181, 1932; Green, R.: "Report on 50 Cases of Malaria Treated With Atebrin," *Lancet 1:* 826, 1932) did not fully answer the question of whether Atabrine was fundamentally superior to quinine, although it has recently been

claimed that Atabrine gives more prompt and effective control of parasitemia and symptoms and longer intervals until relapse (Most, Harry, and Hayman, Joseph M., Jr.: "Relative Efficiency of Quinacrine [Atabrine] and Quinine in Treatment of Acute Attacks of Vivax Malaria," *Am. J. M. Sc. 211:* 320, 1946). Shortage of quinine during World War II led to renewed study of Atabrine. It was shown that its chemotherapeutic activity was a simple function of the plasma concentration ("Plasma Quinacrine Concentration as a Function of Dosage and Environment: A Joint Report of the Armored Medical Research Laboratory, Fort Knox, Kentucky, and the Commission on Tropical Diseases, Army Epidemiological Board, Preventive Medicine Service, Office of the Surgeon General, United States Army," *Arch. Int. Med. 78:* 64, 1946). This work led to more intelligent dosage schedules for suppression and therapy of acute attacks. The failure to prevent relapses (radical cure), the development of more effective drugs, and the occurrence of severe toxic symptoms have led to its displacement. The violent skin reactions are described in a series of articles in the *Bulletin of the U. S. Army Medical Department* (*4:* 653, 687, 724, 725 [Dec.] 1925).

VOL. 1 195

42. Eaton, Paul: "Susceptibility of Red Cells to Malaria," *Am. J. Trop. Med. 14:* 431, 1934.

The question had often been raised of why malarial parasites do not multiply until every blood cell is infected. Eaton offered an explanation of this in a case of inoculation malaria (presumably tertian) when he found that the parasite seemed only to invade the reticulocytes (young cells). He presumed that schizonts which do not infect are promptly phagocytosed. This work was amplified by R. Hegner ("Relative Frequency of Ring-Stage Plasmodia in Reticulocytes and Mature Erythrocytes in Man and Monkey," *Am. J. Hyg. 27:* 690, 1938), who found that tertian parasites mostly invaded young cells, but quartan and falciparum parasites were found mainly in mature red cells. The exact meaning of all this work is not clear. It is of interest that R. Craick ("The Erythrocyte in Malaria," *Lancet 1:* 1110, 1920) had found that stippled red cells resulting from ingestion of lead were prone to invasion by *P. vivax.*

43. Haggis, A. W.: "Fundamental Errors in the Early History of Cinchona," *Bull. Hist. Med. 10:* 417, 1941.

For some 300 years, until the development of quinacrine, cinchona bark or its alkaloids were the only effective therapy for malaria; history owes quinine an unrepayable debt, even though in the last ten years it has been largely replaced by more effective drugs (Refs. 40, 47, 51, 53). Haggis, in an important study, seems to disprove the traditional view that the Countess of Chinchon introduced the remedy into Spain when she returned with her husband from Peru. The full history of the terms "cinchona" and "quina-quina" is discussed.

At the date when this bibliography begins, the use of Peruvian bark in intermittent fever was common knowledge. The subsequent literature is vast.

The quinine alkaloids were extracted by Pelletier and Caventou in 1820 ("Recherches chimiques sur les quinquinas," *Ann. chim. et phys. 15:* 289, 1820) and have been generally used since then. Considerations of the action of quinine important to the clinician are reviewed by Laveran (Ref. 11, 1891, p. 183), by Thayer and Hewetson (Ref. 22), and by Ross (Ref. 36), and a more modern summary is that of Erwin E. Nelson ("Cinchona and Its Alkaloids in the Treatment of Malaria," in *A Symposium on Human Malaria*, Washington, 1941, American Association for the Advancement of Science, Smithsonian Institution Building, p. 255). The realization that quinine, even in large doses, could not be depended on for radical cure of malaria became clear when it was found that it failed to eliminate the extraerythrocytic cycle of the types of malignant vivax infection seen in the South Pacific in World War II (Ref. 46).

44. Temkin, Owsei, and Ramsay, Elizabeth: *Antimalarial Drugs*, Washington, 1944, National Research Council, Division of Medical Sciences. Issued by the Office of Medical Information.

At the beginning of World War II quinine was still the standard remedy for malaria and Atabrine was the only other known drug which had any noteworthy effect. The shortage of the supply of quinine due to enemy action and the realization that neither quinine nor Atabrine was effective in completely eradicating the disease and preventing relapses led to a frantic search for new antimalarials; and thousands of prospective compounds were developed and tested. Out of this work came the new drugs which are effective both in suppression and eradication. An idea of the "antimalarial drugs" situation in 1944 is obtained from this monograph "intended to give the general background of present knowledge of the treatment of malaria by drugs under existing war conditions."

It should be emphasized that the few references on chemotherapy which we are able to include in this bibliography give no idea of the thousands of papers dealing with effects of antimalarials. Authoritative reviews to date are that of J. A. Shannon ("Study of Antimalarials and Antimalarial Activity in Human Malarias," *Harvey Lectures 41:* 43, 1945-1946) and that of N. Hamilton Farley ("Malaria: Lecture II. Chemotherapy," *Brit. M. J. 2:* 891, 1949).

45. Ash, J. E., and Spitz, Sophie: *Pathology of Tropical Diseases: An Atlas*, Philadelphia and London, 1945, W. B. Saunders Company, "Malaria," p. 206.

This is an authoritative account with beautiful plates of the pathology of malaria. For a full description of the microscopic anatomy with extensive bibliography consult Taliaferro, William H., and Mulligan, H. W.: "The Histopathology of Malaria With Special Reference to the Function and Origin of the Macrophages in Defence," *Indian M. Research Mem.* Calcutta, 1933.

46. Fairley, N. Hamilton: "Chemotherapeutic Suppression and Prophylaxis in Malaria," *Tr. Roy. Soc. Trop. Med. & Hyg. 38*: 311, 1945.

The idea of an exoerythrocytic phase in human malaria was "in the air" in the 40's, and Fairley's work brought strong indirect evidence. The failure of such drugs as quinine, which eliminate trophozoite-transmitted (inoculation) malaria, to prevent relapses and produce permanent cure suggests that quinine-resistant forms must be elsewhere in the tissues than in the blood. Fairley also showed by subinoculation experiments that vivax sporozoites injected by mosquito bite disappeared from the blood to reappear nine days later; he presumed that they were developing outside the circulation. Further elaborate and most instructive observations by Fairley and his associates on the results of subinoculations and the effect of drugs under various conditions ("Researches on Paludrine [M. 4888] in Malaria," *ibid. 40:* 105, 1946, and "Sidelights on Malaria in Man Obtained by Subinoculation Experiments," *ibid. 40:* 621, 1947) were made in the next few years; but it remained for Shortt and Garnham (Ref. 50) actually to demonstrate the pre-erythrocytic tissue stage in man.

Fairley also presumed that in contrast to vivax, *P. falciparum* has only a short-lived tissue stage; there is a brief pre-erythrocytic stage but no subsequent exoerythrocytic stage. Radical cure can therefore be achieved by drugs such as quinine, Atabrine, and chloroquine which have only a suppressive effect in vivax malaria. Fairley's important summary of the whole problem should also be consulted (Ref. 52).

The early views on relapses are discussed in Ross's book (Ref. 36, p. 108). Some of these theories were that relapse was due to inadequate quinine therapy or temporary "encystment" of parasites in internal organs whence they "broke out" or that crescents kept infection going. Ross believed that "rallies" were due to extensive destruction of parasites by some immune process; when this flagged, parasites "escaped" and relapse occurred. Many traumatic factors such as cold, fatigue, illness, or excessive alcohol might be precipitating circumstances. C. F. Craig, the great American malariologist ("A Study of Latent and Recurrent Malarial Infection and the Significance of Intracorpuscular Conjugation in the Malarial Plasmodia," *J. Infect. Dis. 4:* 108, 1907) thou ght that "intracorpuscular conjugation" produced a resting stage of the plasmodia resistant to quinine, which later gave birth to young plasmodia, thus causing a relapse. This view was widely accepted. The true explanation of relapses was not fully understood until the exoerythrocytic stage of the parasite in man was worked out (Ref. 50).

47. Most, H., London, I. M., Kane, C. A., Lavietes, P. H., Schroeder, E. F., and Hayman, J. M.: "Chloroquine for Treatment of Acute Attacks of Vivax Malaria," *J. A. M. A. 131:* 963, 1946.

Chloroquine, a 4-amino-quinoline derivative, turned out to be one of the most effective of the new synthetic antimalarials developed during the War. Its properties were described in a statement approved by the Board for Coordination of Malarial Studies ("Activity of a New Antimalarial Agent Chloro-

VOL 1 195

quine [S.N. 7618]," *J. A. M. A. 130:* 1069, 1946). Most and his associates in clinical trials found it superior to quinine and quinacrine in acute attacks of vivax malaria. This finding was generally accepted, but Pullman and his associates soon showed that the relapse rate after treatment with quinine, quinacrine, and chloroquine was about the same, 90 per cent or over (Pullman, T. N., Branch, C., Jr., Alving, A. S., Whorton, C. M., Jones, R., Jr., and Eichelberger, L.: "Comparison of Chloroquine, Quinacrine [Atabrine] and Quinine, in the Treatment of Acute Attacks of Sporozoite-Induced *Vivax* Malaria [Chesson Strain]," *J. Clin. Invest. 27:* No. 3, Part 2: 46, 1948). Hence the drug did not produce radical cure, but it is an important survivor of the vast number of potential antimalarials tested.

48. "Symposium on Biochemistry of Malarial Parasites," *Fed. Proc. 5:* 390, 1946.

OL.
1
955

Although the subject is outside the scope of this bibliography, those interested will find here several papers which deal authoritatively with this part of the subject.

49. MacDougall, Mary Stuart: "Cytological Studies of Plasmodium: The Male Gamete," *J. Nat. Malaria Soc. 6:* 91, 1947.

The significance of the sexual forms gradually became clear through the work of MacCallum (Ref. 24) and others (Ref. 23). It is of interest that chromosomes seem now to have been definitely demonstrated in the various sexual forms. Many interesting and important details are brought out in this paper by MacDougall.

50. Shortt, H. E., and Garnham, P. C. C.: "The Pre-Erythrocytic Development of Plasmodium Cynomolgi and Plasmodium Vivax," *Tr. Roy. Soc. Trop. Med. & Hyg. 41:* 785, 1948.

Although an exoerythrocytic cycle in bird malaria had been described as early as 1936 and in monkey malaria later (for review of the subject see Huff, C. G., and Coulston, F. J.: "The Development of Plasmodium Galenaceum From Sporozoite to Erythrocytic Trophozoite," *J. Infect. Dis. 75:* 231, 1944), the human exoerythrocytic cycle was first reported in a brief note (Shortt, H. E., Garnham, P. C. C., Covell, G., and Shute, P. G.: "The Pre-Erythrocytic Stage of Human Malaria, Plasmodium Vivax," *Brit. M. J. 1:* 547, 1948). The findings are described in more detail in the present paper. It is pointed out that after "inoculation" malaria there is really no incubation period, whereas after infection by mosquito bite there is a "clear cut incubation period varying in length according to the species of parasite." It is also noted that inoculation malaria can be promptly eradicated by drugs, whereas "after sporozoite inoculation [by mosquito] the disease will make its appearance sooner or later in spite of the drug." All this strongly suggested a different sequence of infection with merozoites and sporozoites.

A volunteer was bitten by a huge number of anopheles infected with *P. vivax* and was also given an intravenous injection of an emulsion of infected mosquito salivary glands. Seven days later *liver* biopsy showed "plasmodial masses studded with chromatin particles" which were interpreted as merozoites.

Interesting immunologic implications of the extraerythrocytic cycle are also discussed.

Shortt, Fairley, Covell, Shute, and Garnham ("Pre-Erythrocytic Stages of Plasmodium Falciparum," *ibid. 44:* 405, 1951) later described the pre-erythrocytic stage of *Plasmodium falciparum* in two volunteers. Liver biopsies on the fourth, fifth, and sixth days after inoculation showed beautifully the pre-erythrocytic schizonts in parenchymal cells.

For a review of the pertinent literature leading up to the whole question of the pre-erythrocytic and extraerythrocytic cycle in man see N. Hamilton Fairley's penetrating review (Ref. 52) and P. C. C. Garnham's authoritative summary, with literature ("Exo-Erythrocytic Schizogony in Malaria," *Trop. Dis. Bull. 45:* 831, 1948). For a brief discussion see the recent article by P. C. C. Garnham ("The Life History of the Malaria Parasite," in *Lectures on the Scientific Basis of Medicine 2:* 323, The Athlone Press, 1954). Thus the sequence has been established in vivax malaria of the erythrocytic cycle (in man), the cycle in the mosquito, the pre-erythrocytic cycle in man, erythrocytic cycle again, followed by exoerythrocytic cycle in tissues.

51. Alving, Alf S., Craig, Branch, Jones, Ralph, Jr., Whorton, C. Merrill, Pullman, Theodore N., and Eichelberger, Lillian: "Pentaquine (Sn-13, 276), a Therapeutic Agent Effective in Reducing the Relapse Rate in *Vivax Malaria*," *J. Clin. Invest. 27:* No. 3, Part 2: 25, 1948.

In a search for agents which might have a more potent curative effect, the authors (*ibid.*, p. 34) had conducted clinical trails with eighteen analogues of pamaquine. Pentaquine seemed most promising. The drug was tested in volunteers infected with vivax malaria. While its effect in suppressing fever and parasitemia was not significantly different from quinine alone or quinine and pentaquine administered concurrently, among thirty subjects only one suffered a relapse. Pentaquine was thought to be somewhat more effective than pamaquine in producing radical cure. This work has been confirmed by others, including Coggeshall, and Rice ("Cure of Chronic Vivax Malaria With Pentaquine," *J. A. M. A. 139:* 437, 1949) and Straus and Gennis ("Evaluation of Pentaquine as a Cure of Relapsing Vivax Malaria," *Bull. New York Acad. Med. 24:* 395, 1948; "Radical Cure of Relapsing Vivax Malaria With Pentaquine-Quinine: A Controlled Study," *Ann. Int. Med. 33:* 1413, 1950).

52. Fairley, N. Hamilton: "Malaria, With Special Reference to Certain Experimental, Clinical and Chemotherapeutic Investigations," *Brit. M. J. 2:* 825, 1949.

Innumerable observations on resistance and immunity may be culled from the early writings. King (Ref. 21), for example, noted the resistance of African

Negroes to malaria and thought that the pigmentation which gradually developed in white skins as the disease became chronic was a desirable event which indicated increased resistance. Laveran (Ref. 11, 1891, p. 175) recognized the fact that "cures called spontaneous, that is, without the use of quinine, are not rare." He believed, following Metchinkoff, that phagocytosis of parasites by leukocytes and macrophages was the associated factor in immunity. Koch's remarks (Ref. 32) are also of interest.

The modern views on immunity are summarized by L. T. Coggeshall ("Immunity in Malaria," *Medicine 22:* 87, 1943), but the position of the extraerythrocytic phase in immunity is more fully developed by Fairley in this lecture and by P. C. C. Garnham and H. E. Shortt ("Demonstration of a Persisting Exo-Erythrocytic Cycle in Plasmodium Cynomolgi and Its Bearing on the Production of Relapses," *Brit. M. J. 1:* 1225, 1948). These workers believe that the "antiparasitic defence mechanism" of the host is directed against the erythrocytic cycle and not against pre-erythrocytic or exoerythrocytic forms, which seem to be protected against destruction by their situation in the liver cells. They think there is a "specific cellulo-humoral response" to red cell invasion by the erythrocytic parasites. Clinical immunity does not mean that every parasite is permanently cleared from the blood. The parasites actually seem to be destroyed in various internal organs. The relations of latency and relapse, especially when the situation is modified by therapy, to immunity become therefore excessively complicated. Another interesting phase is specificity of immunity to species or even to strains. This question is discussed by Coggeshall (loc. cit.), and active work, summarized by Young (Ref. 54), is at present going on.

OL.
1
955

53. Gennis, Joseph, Straus, Bernard, Kenney, Michael, and Klein, Bernard: "The Use of Primaquine for the Treatment of Malaria in Korean Veterans," *Am. J. Med. 17:* 223, 1954.

Primaquine, still another quinoline derivative, was recently discussed in a comprehensive report to the Council on Pharmacy and Chemistry ("Status of Primaquine," *J. A. M. A. 149:* 1558, 1952). Gennis and his associates de-

TABLE I

	SUPPRESSION OF SYMPTOMS (PROPHYLAXIS)	CURE OF CLINICAL ATTACK	RADICAL CURE (PREVENTION OF RELAPSE)
Quinine	+	+	0
Quinacrine (Atabrine)	++	++	0
Chloraquine Paludrine	+++	+++	0
Plasmochin (pamaquine)	±	±	+
Pentaquine	±	±	++
Primaquine	±	±	+++

scribed the use of this compound in Korean conflict veterans with vivax malaria. One hundred one patients were treated with only one relapse. There were four recurrences (return of fever in less than one month after therapy). The writers advise, therefore, combined therapy with chloroquine for its suppressive effect with primaquine for radical cure. This seems to be the most effective treatment today for malaria.

Table I illustrates the development of drug therapy as outlined in Refs. 47, 51.

This table is an oversimplification and doubtless will soon be modified by further work as new antimalarials are constantly being tested. Meanwhile the subject of chemotherapy is brilliantly reviewed by N. Hamilton Fairley ("Malaria: Lecture II. Chemotherapy," *Brit. M. J. 2:* 891, 1949).

54. Young, Martin D.: "Malaria During the Last Decade," *Am. J. Trop. Med. & Hyg. 2:* 347, 1953.

This is a review of important recent contributions.

VOL
1
195

Editorials

ON THOSE WHO CARE FOR THE SICK

Recent medical literature, together with the public press, has drawn attention to the problems of chronic disease in our time. The social, economic, and medical aspects of chronic disease have been dealt with in scientific and lay publications and through the mediums of conferences and fund-raising campaigns. The publication of the first issue of the JOURNAL OF CHRONIC DISEASES indicates the need for a medical journal devoted to the problems of chronic diseases of all age groups.

Although the greater number of the scientific contributions to the understanding of chronic disease have been concerned primarily, if not exclusively, with the nature and course of disease and disability, it is both interesting and encouraging to note an increasing trend toward systematic inquiry into the personal professional experiences of those who study, care for, and are responsible for the chronically sick. If not unique, these experiences are different from the experiences of those who care for the acutely ill.

We should remind ourselves that the systematic and objective study of the physician-patient relationship is very recent. As with other human relationships heretofore considered in intuitive and artistic terms, it is to be expected that progress will be slow and will continue to meet many resistances. It is also well known, even though there are beginnings, that our existing knowledge of the more significant human relationships between the physician and others who study and care for the sick is limited and tentative. For example, studies of the dynamic relationships of nurses, aids, attendants, and physicians, with each other, with their patients and their families in institutions and in patients' homes, either do not exist or have just been begun. Knowledge of these matters, limited as it is in the total field of human illness, is even more limited in chronic disease.

Currently, the tradition of medicine in the medical school as well as in the public eye through the modern novel, drama, and moving picture centers principally about the immediate life-death problems of acute illness, of injury, of diagnosis, of dramatic critical decisions. We learn from Sigerist[1] that throughout history the person who is chronically sick or disabled has been treated differently from the person who recovers quickly from illness. In primitive tribes, in ancient Semitic civilization, in classical Greek society, human illness, however understood in concepts of causes and effects, was regarded in terms of socially deviant behavior, and thus the sick and disabled could expect consideration from society only so long as their conditions were capable of improvement. That many of these attitudes have persisted long after the introduction of Christianity and into the modern period of scientific or naturalistic attitude toward

illness is attested by current attitudes toward the physically disabled and deformed and more strikingly by attitudes toward the epileptic, the mentally defective, and those who are chronically mentally sick.

Howard Rusk's[2] recent statement that changes must take place in the medical school to enable the physician of tomorrow to become more interested and better prepared to understand, study, and care for the chronically sick person is of fundamental significance. The experiences of most students in our teaching hospitals and their outpatient departments do not give them an adequate opportunity for long-term study and continuity of work with patients. This is not only the result of hospital administrative structure and of the real differences of curricular planning of the medical student teaching program, but is also a matter of a lack of a model. Medical students learn not only through precept but by example. Few students have the opportunity to work with and thus to be identified with skilled, seasoned, stimulating medical teachers and investigators responsible for and engaged in the study and care of the chronically sick. Students' interest in the study, care, and treatment of the chronically sick must be determined in great part by their awareness and recognition of the problem and their opportunity to work and identify with those who are devoted and dedicated to these problems. The continued responsibility for the care of a chronically sick person adds immeasurably to the education of a physician. Moreover, it provides indispensable preparation for the work he is to do as a practicing physician. There are interesting and exciting beginnings of experiments in undergraduate medical education. There has been an increase in preceptorship programs in a number of medical schools. Some schools have introduced family and home care programs. Other schools have begun to assign students regularly to patients who are chronically sick. It will be important to learn from these experiments ways and means to promote and to further the opportunity for the medical student to study the chronically sick.

The emotional experiences encountered among those who care for the chronically sick are different from those experiences with the acutely ill. Contrast the gratifying satisfactions of those who understand clearly the genesis and current state of their patient's disability or their skill in applying successfully existing knowledge leading to the amelioration of symptoms or to complete recovery of the patient with the continuing responsibility of care for patients who continue to experience pain, distress, or disability. In the latter instance there may be limited knowledge of the disabling process and of its course. Current knowledge and skill accumulated painstakingly with great effort and sacrifice when applied to the care of the patient do not necessarily relieve his discomfort or hasten his recovery. While this is challenging, it is also frustrating and may lead to resentment. The therapeutic intent of the physician, his wish to heal, to cure, to do something positive for the patient, is blocked. When one is responsible for the daily or regular care of such a patient, it is not unexpected that one may become discouraged and, unwittingly, resentful and discouraging. It requires considerable emotional maturity to care, day in and day out, for those who remain disabled or for those who recover slowly and undramatically over a period of months and years.

It is in this setting that the physician must learn to master his own anxieties and his own resentments so that he can continue to help his patient and to develop skills of simple supportive psychotherapy.

Those who care for the chronically sick find that they are concerned not only with the disease, but increasingly with the patient and his disability, with the members of the patient's family, and with the community to which the patient and his family belong. The physician learns that he will need help from the members of the patient's family and from the many health, welfare, and social agencies in the community which exist to serve him in the care of his patient. In chronic illness there are likely to be more serious dislocations of family relationships in their economic and financial aspects. In chronic illness more than in acute illness the physician needs the skills of other members of the health team including the social case worker, rehabilitation experts, physiotherapists, and visiting nurses. It is through chronic illness that one understands why the family has been defined as an autocracy ruled by its sickest member. Care of the chronically ill demands of the physician skills to provide constant and open communication between the patient, the nurse, and himself and with the members of the patient's family. The physician will not only gain inestimably in his understanding of his patient from the different perceptions of others who care for the patient. He also plays an important and indispensable role in fostering and maintaining the morale of the patient and his family. Inevitable human emotions, fear, perplexity, anger, shame, and guilt, are experienced not only by the patient but by the members of his family. It is important that the physician have some basic understanding of the psychology of chronic illness, of prolonged convalescence and disability so that his mind can be freed so far as possible from enslaving emotions in order that he can deal clearly and intelligently with the emotional problems of those whom he is attempting to help.

JOHN ROMANO

REFERENCES

1. Sigerist, H. E.: Civilization and Disease, Ithaca, 1945, Cornell University Press.
2. Rusk, H.: Chronic Illness and the Medicine of Tomorrow, J. CHRON. DIS. 1:88 (Jan.) 1955.

MANAGEMENT OF THE PATIENT WITH HODGKIN'S DISEASE

Seven per cent of all malignant neoplastic diseases are due to the lymphomas and leukemias, and the tumor described by Thomas Hodgkin accounts for one-sixth of this group. In New York State the incidence of Hodgkin's disease is 2.5 per 100,000 population, a frequency comparable to carcinoma of the tongue in men or carcinoma of the kidney in women.[1] The disease affects each sex equally, and although the incidence is highest in the age range of 70 to 85, the disease is most frequently seen in young adults between the ages of 25 and 50. The availability of new and improved methods of treatment, the clarification of concepts of management, and the importance of planning and interspecialty cooperation for the best result to the patient justify a consideration and evaluation of therapy in Hodgkin's disease at this time.

In order to devise the best regimen with the therapeutic measures currently available, a knowledge of the natural history of Hodgkin's disease is essential. For the purposes of the present discussion four important stages will be considered. In an indeterminate proportion of all the cases the tumor arises in a single focus, which usually is a superficial lymph node or group of lymph nodes. This primary stage of the disease does not occur in all instances, or perhaps it goes unrecognized by the patient and his physician. The phase of localized tumor is followed by regional involvement, and not infrequently this second stage of dissemination merges imperceptibly into obviously generalized disease. In the first two stages of Hodgkin's disease, symptoms are characteristically absent and the physical finding of lymphadenopathy by palpation or x-ray examination offers the only means of detection. As the tumor becomes more widespread, pressure by enlarged nodes or infiltrated viscera on blood vessels, nerves, or other adjacent structures leads to mechanical disturbances in function. More devastating than these, and less well understood, are the modifications in homeostatic mechanisms and normal cell functions reflected by fever, hemolytic anemia, abnormal protein synthesis, pruritus, and anorexia. The final stage of Hodgkin's disease is characterized by intensification and progression of the features of the preceding phase in which, in addition, pancytopenia, with its accompanying complications, and cachexia are prominent.

The early diagnosis of Hodgkin's disease rests chiefly upon careful physical examination followed by lymph node biopsy and histopathologic study. It is to be emphasized that clinical diagnosis unconfirmed by microscopic tissue examination jeopardizes the planning and implementation of an orderly, effective, and long-term therapeutic schedule. The histologic classification of Hodgkin's disease into paragranuloma, granuloma, and sarcoma appears to have little prognostic importance and certainly does not alter the type or sequence of treatment.

VOL.
1
195

When Hodgkin's disease is localized, aggressive local therapy is indicated. It has been repeatedly demonstrated in series reported by such competent authorities as Craver,[2] Holmes and Schulz,[3] Lenz,[4] Stout,[5] and Peters[6] that vigorous treatment of Hodgkin's disease when, clinically, only one site is involved leads to five-year survival in 25 to 40 per cent of the cases. This figure drops precipitately to less than 10 per cent if three or more sites are involved when treatment is first initiated. Needless to say, the majority of patients with apparently localized Hodgkin's disease unfortunately have more disseminated disease than is clinically recognized. Since the physician can learn this only by watching the evolution of the natural history of the disease, the patient must be given the opportunity for cure even though the maximal effort may be doomed from the start. The type of radical curative therapy for apparently localized Hodgkin's disease is either surgery or radiotherapy. Extensive experience with the surgical attack on this tumor has not been recorded, and even though small series of carefully selected patients indicate good results,[7] the conventional approach to the problem will probably continue to be by x-irradiation. It would be presumptuous for me to present the details of radiotherapy. It nevertheless is important for the internist to know that aggressive radiotherapy implies a tumor dose of 3,000 r. to 5,000 r., which far exceeds the amount required to cause disappearance of palpable tumor. A patient's chance for long survival may be jeopardized by the application of a small dose which decreases the size of the radiosensitive tumor but fails to destroy it completely.

Before leaving the treatment of localized Hodgkin's disease, there is one additional problem which should be mentioned. Not infrequently the primary manifestation of a malignant lymphoma is in hilar or mediastinal lymph nodes. In the absence of peripheral lymphadenopathy the question inevitably arises whether or not the patient should be subjected to an exploratory thoracotomy for diagnostic purposes. The alternative which is usually proposed, and most frequently accepted, is presumptive diagnosis by a therapeutic trial of external radiation. It is to be emphasized that this diagnostic technique does no more than indicate a radiosensitive or radioresistant tumor. Since it is not unusual for a lymphoma to be radioresistant and an epithelioma to be radiosensitive, it is obvious that the therapeutic trial does little to define the problem precisely. Since the techniques of thoracic surgery and anesthesia have been perfected to the point where the operative risk of opening the chest is negligible, a definitive diagnosis by formal biopsy should be very seriously considered. The direct visualization of the tumor provides the opportunity not only for histopathologic documentation, but also for the removal of a benign tumor which roentgenographically may be indistinguishable from a malignant neoplasm or the outlining by silver clips of the limits of a malignancy if this is found, so that the radiotherapeutic attack may be more effective.

When Hodgkin's disease is no longer localized to a single group of lymph glands but involves nodes in one region, vigorous radiotherapy is still indicated. The definition of a region would include unilateral cervical nodes and the mediastinum; cervical and axillary lymph nodes on one side; or unilateral femoral and iliac lymphadenopathy. When this stage of the disease has been reached,

one can anticipate that tumor extension will occur in apparent bursts of activity which recur with increasing frequency until there is widespread anatomic involvement. The biologic characteristics of the neoplasm and, perhaps to a lesser extent, therapy determine the rapidity with which dissemination occurs. Certainly it has been observed repeatedly that aggressive radiotherapy of regional disease can be succeeded by long remission intervals.

When the disease has become generalized and symptomatic, the aim of treatment is palliation. This is an accurate statement of the objectives of therapy for widespread Hodgkin's disease, but the word "palliation" too often is interpreted to mean sedation plus sympathetic neglect. The palliative management of disseminated malignant lymphomas implies an active program which, if administered wisely, can provide a significant prolongation of asymptomatic and useful life. In this stage of Hodgkin's disease the constitutional symptoms mentioned earlier may be very intense, with a swinging temperature curve up to 103° to 105° F., night sweats, severe malaise, and anorexia. The physical examination may be unimpressive, with little or no superficial lymphadenopathy, and minimal hepatosplenomegaly. Under such circumstances it can be reasonably presumed that there is significant retroperitoneal tumor usually distributed in the para-aortic nodes. Before the advent of chemotherapeutic agents radiotherapy was directed to this region with successful production of remissions in the majority of cases.

At the present time, the intravenous injection of nitrogen mustard can replace radiotherapy for the treatment of the patient acutely ill with Hodgkin's disease. This drug offers the advantages of rapid onset of effect so that improvement can be anticipated within three to five days after the completion of the nitrogen mustard administration. The time-tested total dosage of the compound for each course is 0.4 mg. per kilogram of body weight, and this may be conveniently given in two doses separated by a twenty-four-hour interval. Courses of therapy should be spaced at intervals of at least one month and preferably longer. The most distressing toxicity following the intravenous administration of nitrogen mustard is nausea and vomiting, which occur from one-half hour to two hours after drug administration and may persist with diminishing severity for two to five days after the last injection. This very unpleasant feature of therapy has been largely eliminated by the use of chlorpromazine, which may be given in doses of 25 to 50 mg. by mouth four times daily starting several hours before the first injection and continuing for at least one day after the completion of the course.

The important toxicity of nitrogen mustard, or any of its congeners, is depression of bone marrow function. For this reason the drug is contraindicated in the presence of leukopenia or thrombocytopenia because these will inevitably be accentuated by the treatment. There is one exception to this rule: Occasionally pancytopenia is a prominent feature in an untreated patient. When the evidence strongly indicates that this is a primary manifestation of Hodgkin's disease rather than secondary to previous treatment, nitrogen mustard therapy may be followed by a very satisfactory improvement in the peripheral blood counts. The mechanism for the pancytopenia just mentioned is hypersplenism.

VOL 1 195

Splenectomy has been performed in a number of patients with this functional disorder complicating Hodgkin's disease. Although hematologic improvement was noted by both Dameshek[8] and Karnofsky[9] in the majority of the patients, this was transient and the fundamental course of the disease was not altered. Anemia is a very prominent finding in the group of patients with constitutional manifestations. It has been well established by Ross[10] and by Hyman[11] that this is due to an increased rate of hemolysis. In these individuals the lifespan of transfused erythrocytes may be as short as twenty days and therefore blood replacement is of little benefit. It is preferable, therefore, to attempt to induce a remission with nitrogen mustard before transfusions, since clinical experience has shown that the hemoglobin level will be better maintained at this time.

In 1951 a critical evaluation of the contribution of nitrogen mustard in the therapeutic management of Hodgkin's disease was published based on the experience at the Columbia-Presbyterian Medical Center. The course of the disease was compared in two groups of patients: sixty-seven who received radiation and nitrogen mustard in alternating therapeutic courses as indicated, and sixty-five who received radiotherapy alone. The available evidence showed that the survival rate from the onset of symptoms, from the time of biopsy diagnosis, or from the time of evident generalized disease was not significantly different in the two groups of patients. Although this led to the conclusion that the addition of nitrogen mustard had no demonstrable effect in modifying the duration or life expectancy in the disease, evidence was presented which showed that the amount of time necessary for radiation treatments was reduced by one-half, the asymptomatic period of life increased proportionately, and the economic burden was made lighter.[12]

This discussion of the management of generalized Hodgkin's disease has been concerned with the acutely ill patient. Frequently the disease activity may be characterized as smouldering with correspondingly less intense signs and symptoms. The patient can be ambulatory and his chemotherapeutic treatment can also be modified. The most useful drug under this circumstance is triethylene melamine (TEM), which is an orally active congener of nitrogen mustard.[13] Since it undergoes chemical reaction in acid media, the absorption of TEM from the gastrointestinal tract can be made more uniform by administration together with 2 Gm. of sodium bicarbonate after an overnight fast.[14] Breakfast is then withheld for an additional two to four hours. In an adult, without hematopoietic depression, 5 mg. given on three successive days is a reasonable dose. This should be reduced for persons under 55 kilograms, and for infants and children the total dose is 3 to 5 mg. Nausea and vomiting are less frequent and less severe than with intravenous nitrogen mustard. The onset of gastrointestinal symptomatology is usually delayed for six to ten hours after drug administration and can be controlled with chlorpromazine. Therapeutic benefit from triethylene melamine is slow in appearance and ten days to two weeks may pass before it is evident. Similarly the evidence of bone marrow depression may be delayed up to three weeks, and failure to observe depression of the circulating blood corpuscles before this time should not lead to premature readministration of the drug.

Other compounds chemically related to nitrogen mustard are also available, but these have not been studied extensively nor do they appear to have any attributes which make them obviously superior to nitrogen mustard and triethylene melamine. These include triethylenephosphoramide (TEPA), triethylenethiophosphoramide (ThioTEPA), and p-carboxyproplyphenyl bis beta-chloroethylamine (CB 1348).

The anti-inflammatory actions of cortisone can also be utilized in the management of patients with Hodgkin's disease under special circumstances. Although the evidence for a significant antitumor effect of adrenal cortical hormone therapy is unconvincing, very satisfactory suppression of the fever, malaise, anorexia, and pruritus and, on occasion, decrease in size of peripheral nodes or visceral enlargement can be achieved by cortisone administration. The drug must be given in large doses, 100 to 300 mg. daily, and must be continued without interruption. With time, the signs and symptoms of active disease recur in spite of the hormone therapy. The indication for cortisone, since it modifies symptoms rather than the tumor, is symptomatic generalized disease with pancytopenia, which prohibits therapy with radiation or drugs. Under these circumstances the hormone may provide the time necessary for the return of bone marrow function so that more definitive therapy may be reinstituted. Another indication for cortisone therapy in Hodgkin's disease is the occasional patient whose major manifestation of generalized disease is hemolytic anemia. Usually, hemolytic anemia is only one of many problems of generalized disease, and then nitrogen mustard rather than cortisone is indicated.

The untoward reactions to cortisone are of concern when the drug is used in Hodgkin's disease, since the dosage is necessarily high. These include salt and water retention, dissemination of tuberculosis or other infections, and peptic ulceration. In an attempt to avoid some of these undesirable side effects of cortisone, Butazolidin (phenylbutazone) has been tried.[15] The favorable results with this drug are somewhat less uniform and the salt and water retention is as great as with cortisone. Hydrocortisone offers no significant advantage and there has been insufficient experience to learn whether metacortin will be superior to Compound E.

Brief mention is warranted of several special problems which may appear during the course of the illness. Primary involvement of the central nervous system by this tumor is very rare; however, impairment of function by extradural pressure is not uncommon. Paraplegia and disturbances of bladder and bowel function may be the first signs and symptoms of the disease, requiring laminectomy both for decompression and diagnosis. After this has been established, local radiotherapy is the treatment of choice. If these same manifestations occur during the course of documented Hodgkin's disease, the presumptive diagnosis of extradural tumor is usually accepted and nonsurgical procedures attempted. Because rapid decrease in pressure is so very important for the preservation of cord function, intravenous nitrogen mustard is recommended to achieve quick reduction in the size of the tumor mass. This is followed by local radiotherapy. Such a combined regimen has the advantage over radiotherapy alone not only of decreasing spinal cord pressure rapidly, but also of

avoiding the hazards of an initial increase of pressure secondary to the radiation reaction in the tumor. This latter transient period of edema in and around the tumor can be largely obviated by the administration of small doses of ionizing radiation at the outset, although at the cost of precious time.

Another associated central nervous system complication of Hodgkin's disease and other malignant lymphomas is cryptococcus or torula meningitis.[16] Available evidence suggests that the general debility caused by the malignant lymphomas provides an especially favorable environment for the invasion of this pathogenic fungus, and its presence should be suspected when signs and symptoms of meningeal irritation appear during the course of these diseases. Unfortunately the antifungal drugs now in use have very limited usefulness in the treatment of cryptococcosis.

Over 70 per cent of patients with Hodgkin's disease have intrathoracic involvement at some time during the natural history of their disease. This may present itself as infiltration of hilar or mediastinal lymph nodes and adjacent structures, as discrete or infiltrative lesions of the lung parenchyma, as pleural metastases, or, rarely, as ulcerating lesions of the tracheobronchial passages themselves. The general principles of treatment already discussed apply to Hodgkin's disease of the chest contents, but there are two special circumstances to be mentioned. The superior mediastinal compression syndrome as a complication of tumor involvement of the mediastinum is not infrequent. The increased swelling which may occur as an initial reaction of the tissues to radiation, and the importance of rapid relief from obstruction of the great vessels or trachea, provide the indication for intravenous nitrogen mustard followed by radiotherapy after subjective and objective improvement has been obtained. Intractable effusion due to extension of the tumor to the visceral or parietal pleura may present a distressing problem. If radiotherapy to the hemithorax is contraindicated, and when the intractability of the effusion has been demonstrated by frequent, necessary taps, effective control of the fluid accumulation can be achieved by the intrapleural administration of nitrogen mustard. The injection is given at the completion of thoracentesis, and the recommended dosage is 0.3 mg. per kilogram of body weight. This may or may not cause nausea and vomiting, depending on the rates of fluid exchange across the pleural membrane between the systemic circulation and the pleural cavity.[17,18]

Hodgkin's disease involving the gastrointestinal tract may occur as a primary manifestation of the tumor or as a part of the generalized dissemination of the disease. In the former situation the diagnosis is usually made only after exploratory celiotomy and resection of the involved segment, which may be stomach, small intestine, or colon. Surgical resection followed by abdominal radiotherapy for primary gastrointestinal Hodgkin's disease frequently is followed by long remission intervals[19,20]; the prognosis in this form of the disease is more favorable than in the more usual pattern of primary superficial lymphadenopathy.

A difficult issue is presented to the physician when pregnancy occurs in a patient with Hodgkin's disease. Should he recommend interruption of gestation and if so, why? What are the hazards to the fetus of treatment of active disease? Apodictic statements on these problems are not warranted because of the wide

range of variability among the many variables involved. Although the majority of reports in the literature indicate that pregnancy fails to modify the natural history of Hodgkin's disease,[21] it does not necessarily follow from this that pregnancy should not be terminated. Statistically, the mean life expectancy of a patient with Hodgkin's disease is about three years. It is apparent that, even if pregnancy does not accelerate the course of Hodgkin's disease, the child born has a very high likelihood of losing his mother early in his life. Unless there are some very compelling reasons to the contrary, the known sociologic problems of a motherless child suggest that interruption of an early pregnancy should be seriously considered. For these same reasons, advice to avoid pregnancy is indicated. If, for any of a variety of reasons, the patient has elected to continue the pregnancy and the need for treatment arises, various factors must be noted. Extrapolating from experiments in animals, the hazard to the fetus from radiation or from chemotherapeutic agents is greatest in the first trimester and least in the third trimester. For this reason abdominal radiotherapy and nitrogen mustard or triethylene melamine should be avoided in early pregnancy. Radiotherapy to superficial nodes in the neck or axillae has been given without harm, and this is probably true also of the treatment of intrathoracic disease. If treatment with nitrogen mustard or abdominal irradiation is indicated in the first few months of pregnancy because of constitutional symptoms, then there is a clear medical indication for therapeutic abortion. In the latter half of pregnancy, if treatment for generalized disease is indicated, limited evidence indicates that nitrogen mustard does not lead to fetal death or developmental abnormalities.

This completes a brief survey of specific treatment in Hodgkin's disease. Supportive therapy in the form of antibiotics for intercurrent infections, transfusions for anemia, and adequate diet play an important part in the total care of the patient. And, finally, it cannot be overemphasized that a patient with Hodgkin's disease needs not only a doctor who knows what to do and when to do it but also an understanding human being who can help him to make the adjustments dictated by the many and complex developments during the course of his illness.

ALFRED GELLHORN

VOL. 1 1955

REFERENCES

1. Kraus, A. S.: Annual Average Incidence Cancer Rates, New York State, Exclusive of New New York City, 1949-1951, Department of Health, State of New York.
2. Craver, L. F.: Value of Early Diagnosis of Malignant Lymphomas and Leukemias, Monograph, American Cancer Society, 1952.
3. Holmes, G. W., and Schulz, M. D.: Therapeutic Radiology, Philadelphia, 1950, Lea & Febiger.
4. Lenz, M., Wells, J., and Stout, A. P.: Hodgkin's Disease and Lymphosarcoma in Portmann, U. V.: Clinical Therapeutic Radiology, New York, 1950, Thos. Nelson & Sons.
5. Stout, A. P.: Lymphosarcoma and Hodgkin's Disease, Rhode Island M. J. **32**:436, 1949.
6. Peters, M. V.: A Study of Survivals in Hodgkin's Disease Treated Radiologically, Am. J. Roentgenol. **63**:295, 1950.
7. Williams, R. D., Andrews, N. C., and Zanes, R. P., Jr.: Major Surgery in Hodgkin's Disease, Surg., Gynec. & Obst. **93**:636, 1951.
8. Fisher, J. H., Welch, C. S., and Dameshek, W.: Splenectomy in Leukemia and Leukosarcoma, New England J. Med. **246**:477, 1952.
9. Sykes, M. T., Karnofsky, D. A., McNeer, G. C., and Craver, L. F.: Splenectomy in Far Advanced Hodgkin's Disease. Report of Five Cases, Blood **9**:824, 1954.

10. Ross, J. F.: The Mechanism of Anemia in Leukemia and Malignant Lymphomas, J. Clin. Invest. **31**:668, 1951.
11. Hyman, G. A.: Studies on Anemia of Disseminated Malignant Neoplastic Disease. I. The Hemolytic Factor, Blood **9**:911, 1954.
12. Gellhorn, A., and Collins, V. P.: A Quantitative Evaluation of the Contribution of Nitrogen Mustard to the Therapeutic Management of Hodgkin's Disease, Ann. Int. Med. **35**:1250, 1951.
13. Karnofsky, D. A., Burchenal, J. H., Armistead, G. C., Southam, C. M., Bernstein, J. L., Craver, L. F., and Rhoads, C. P.: Triethylene Melamine in the Treatment of Neoplastic Disease, Arch. Int. Med. **87**:477, 1951.
14. Gellhorn, A., Kligerman, M., and Jaffe, I.: Triethylene Melamine in Clinical Cancer Chemotherapy, Am. J. Med. **13**:428, 1952.
15. Rottino, A., Joffe, A., and Hoffmann, G.: Phenylbutazone (Butazolidin), a Useful Chemotherapeutic Agent for Hodgkin's Disease, Arch. Int. Med. **93**:561, 1954.
16. Collins, V. P., Gellhorn, A., and Trimble, R., Jr.: The Coincidence of Cryptococcosis and Disease of the Reticuloendothelial and Lymphatic Systems, Cancer **4**:883, 1951.
17. Administration and Uses of Radioactive Colloidal Gold 198, Abbott Laboratories, Dept. Radioactive Pharm., June, 1953.
18. Gellhorn, A., and Holland, J.: Neoplastic Diseases, Ann. Rev. Med. **5**:183, 1954.
19. Morris, W. R.: Gastric Hodgkin's Disease, J.A.M.A. **149**:1460, 1952.
20. Craver, L. F.: Some Aspects of the Treatment of Hodgkin's Disease, Cancer **7**:927, 1954.
21. Hultberg, S.: Pregnancy in Hodgkin's Disease, Acta radiol. **41**:277, 1954.

)L.
955

SURGICAL MANAGEMENT OF CHRONIC ULCERATIVE COLITIS

Thirty years ago chronic ulcerative colitis was little more than a descriptive pathologic term. Today it is the name of a well-known clinical syndrome, by no means rare in its occurrence, that still presents unsolved problems. Its etiology is unknown. Numerous suggestions as to its cause and its basic natural history have been advanced, but so far none of these ideas has been generally accepted because of failure of definite proof. It is really not at all certain that the term "chronic ulcerative colitis" may not be inclusive of more than one similar but actually different disease entities. The clinical course of the condition varies greatly from patient to patient, and from time to time in the same patient. Among the uncertainties so characteristic of the subject, as is true of most disorders of unknown origin, is the desirable line of therapeutic attack. In spite of this, it is agreed that medical management is generally the proper initial method of handling patients with chronic ulcerative colitis, and that most cases of the disease never require surgical intervention. It is also generally agreed that some patients, not responding to medical care, may become candidates for operative treatment. Here agreement ceases. The questions of when operation becomes advisable, what form it should take, and how extensive it should be are still not answered to the satisfaction of all those concerned with these problems. The multiplicity of answers indicates that definite settlement of these problems is still to be sought. This situation required me to emphasize the fact that what I present herewith is merely one person's view and should not be taken as authoritative.

The first issue arises as to when one should turn from further dependence on medical measures and resort to surgery. In general, two situations justify this change: the protracted failure of improvement or actual deterioration under long-continued medical care, and the initial fulminating nature of the case that rapidly threatens life itself. The first situation is by no means uncommon. Although it is true that most of the mild or moderate manifestations of chronic ulcerative colitis show improvement, remission, or even cure with medical management, a substantial percentage do not so respond. After weeks or months it becomes apparent that no ground is being gained. Instead there is often a slow increase in all the symptoms and complaints. The patient has become a chronic invalid, and the fear arises that unless the downward progress is checked a fatal outcome threatens. It is true that sometimes such patients, for no clearly understood reason, may abruptly improve without any change having been made in treatment, but this happy event is too rare to warrant hope of its occurrence, and too often death does result in these persistently refractory cases. Just how long to wait before invoking surgical treatment is a matter that cannot be settled by rule of thumb. It depends largely upon the personal views of the physician in charge and of the surgeon who may be

VOL.
1
1955

consulted. I feel that certain general conditions warrant a change in the policy of management. If the patient is ill enough to interfere seriously with ordinary life and function—due to frequent diarrhea, abdominal pain, low fever, loss of weight, blood and pus in stools, anemia, and leukocytosis—and if conscientious and well-planned medical care over four or five months has not resulted in improvement, then I believe that operation is not only justifiable but wise. If the same symptoms and general picture not only fail to improve under treatment, but actually grow worse, the period of trial should be correspondingly shortened, and six or seven weeks may be as long as it is safe to defer active intervention.

The fulminatingly acute cases fall into two groups. One of these is characterized by predominant and alarming hemorrhage from the bowel, associated with a violently hyperemic and friable mucosa of the colon. I recall such a case in a fourth-year medical student. Within a short time of the onset of diarrhea and crampy abdominal pain, he was passing frequent melena of bright red blood in large amounts. In spite of repeated transfusions of whole blood the patient was rapidly losing ground. On the third day of this alarming situation an ileostomy was performed. There was almost immediate cessation of bowel bleeding. The symptomatic recovery was soon complete. After several years the ileostomy was closed. The patient has been seen from time to time, and now after twenty-five years has had no further symptoms of his old trouble. Reference will be made later to this case under discussion of types of surgical treatment.

The other form of fulminating colitis is evidenced not by particularly noteworthy bleeding, but by severe and progressive inflammatory processes in the colon. There is an excessive, almost continuous, diarrhea without much blood but with a great deal of mucopurulent discharge. The fever runs high, the pulse is rapid, the abdomen is very tender and rigid, and there is much pain. I saw such a case in a young woman who had been ill less than three weeks, and who had failed completely to be helped by medical treatment. Operation was advised but refused. Two weeks afterward the patient was admitted to the hospital in moribund condition and died about two hours later. An autopsy revealed a colon riddled with innumerable ulcers, five of which in widely scattered areas had already perforated ante mortem with a resulting widespread peritonitis. Whether operation at the time of first seeing the patient, with defunctioning of the colon by ileostomy, would have brought about a happier result can never be proved, but at least the result could have been no worse.

As was mentioned previously, there is wide agreement in the general principle of surgical intervention in the types of cases just discussed, although differences will of course always exist as to the precise moment when this becomes wise in any given instance. The divergence of views is rather more basic among surgeons concerning the principle to be followed and the actual operative procedures to be employed when the time for surgical treatment has come. A brief historical sketch of the development of surgical handling of chronic ulcerative colitis will make the current situation more understandable. The first use of surgery in resistant cases of this disorder was designed to make more

effective the irrigation of the bowel, which then was a feature of medical treatment. For this purpose it was proposed to irrigate the colon from above downward, in an isoperistaltic direction, by making a cecostomy or appendicostomy with an indwelling tube to admit the irrigating fluid into the cecum. The results of this method were short of being satisfactory. The next surgical attempt was to combine irrigation from above with diversion of the fecal steam, so as to remove the stimulus to peristalsis of intestinal contents in the lumen of the colon. This was accomplished by bringing a loop of low ileum out on the abdominal wall and fixing it there. The loop was cut across, and the upper stump formed an ileostomy to pour out the intestinal contents externally, while the lower stump served to admit irrigating fluids to the colon from above. To many surgeons this device seemed to contain an element of self-contradiction. Why, they argued, should one make an ileostomy to put the colon at rest, and then proceed to irritate and stimulate it by the frequent introduction of large quantities of fluid? The real purpose of the ileostomy was not to provide ingress for irrigating fluids of very doubtful utility, but to provide egress of intestinal contents to put the colon at physiologic rest. Under this treatment the colon might heal, just as an infected joint might heal when put in a plaster casing. Accordingly many surgeons modified the operation. The loop of low ileum was cut completely across. The lower stump was closed and replaced in the abdomen. The upper stump was brought out of the incision and fixed to the abdominal wall to form a terminal ileostomy of a semipermanent nature.

The term "semipermanent" is used in view of the varying results that were found to follow this type of operation. These results fell into three groups. The first, the most desirable, and the least numerous group is that in which the diseased colon heals. All symptoms disappear. The proctoscopic and radiologic examinations show a return to normal. Blood and pus can no longer be found in the washings of the bowel. After waiting an adequate time to feel reasonably safe against recurrence—and this means many months and even then is not entirely safe—the ileostomy may be closed and the continuity of the alimentary tract restored. The case of the medical student mentioned earlier in this paper falls in this group. I have had a small number of similar cases.

The second group, much larger, comprises patients whose symptoms practically disappear but whose colons continue to show evidences of disease. The proctoscope reveals a mucosa still hyperemic, that bleeds easily when scraped with an instrument. There are still red and white corpuscles in the lumen of the bowel. These are the cases of which the late Dr. Dan Jones, of Boston, used to say, "The patient is cured, but his bowel is not." Such patients would be in great danger of reactivation of the disease if the ileostomy were closed, and are probably doomed to live with it permanently.

The third group may be described briefly as those in whom the ileostomy is a therapeutic failure. The downward course of the disease continues, not appreciably retarded by the ileostomy. For these people, colectomy is now the operation that offers almost the only chance of recovery. The word "almost" is inserted advisably, because of the unpredictable and very improbable chance of a spontaneous recession of the disease. Under conditions of modern surgery, colectomy either total or subtotal, and performed either in one or two stages,

VOL
1
195

is a standard operation with a reasonably low mortality. The disease is nearly always completely eradicated except for those unusual instances where the pathologic process extends to the ileum. Patients gain weight, recover good general health, can work, and are comfortable. The management of the ileostomy varies with the individual, but generally can be so successfully achieved that annoyance and embarrassment are minimal.

The fact that total colectomy is now being performed with such generally satisfactory results has led to a further development of the surgical theory of treatment. It is argued, by many potent voices, that since Group I described here—the cures by ileostomy alone—is so small, Group II so numerous and still in jeopardy from a retained unhealed colon, and Group III due for a colectomy anyhow, therefore, the standard surgical attack in cases that require operation should be immediate ileostomy and total or subtotal colectomy.

OL.
1
955

I do not share the view and do not go along with this most recent evolution of the surgical management of ulcerative colitis. I feel that there is no reason to jump to the final and irreversible sacrifice of the colon in all cases not responding to medical care. This would mean the penalizing unnecessarily of the Group I patients, those who recover entirely with ileostomy alone. Granted that the number of these patients is small and that most people do pretty well with an ileostomy, I, for one, would not lightly exchange the normal function of the colon for an ileostomy, and I feel that my patients are equally entitled to defer that decision until it is clearly necessary. The same reasoning applies, with greater numerical frequency but somewhat less pathologic justification, to the Group II cases. Why not let them retain their hope of a restitution to normal even though the quiescent colon is not healed—as long as it remains quiescent. In other words, I hold that ileostomy alone should be the first step of surgical intervention when medicine has failed, and that the further employment of colectomy should be withheld until the effect of ileostomy is determined. There may be a few fulminant cases with the colon so hopelessly damaged that the drastic and dangerous step of immediate colectomy offers the only slight hope and as such it is justified. But this does not apply to the great majority of patients needing surgical help. If and when ileostomy has been established, and the patient evidently is not doing well in spite of this, the colon can be removed in a well-planned and unhurried manner.

There is another psychologic reason for adopting the principle just enunciated. Most surgeons, regardless of their detailed operative methods, believe that many cases are withheld from surgical care longer than their best interests require. I am of this opinion myself. However, I feel sure that family doctors, medical consultants, and most of all the patient would be ready sooner and more willingly to accept the suggestion of surgical intervention, if such intervention were initially limited to ileostomy, with the question of colectomy deferred for later consideration. With this as the standard approach, many patients might be given the benefit of the arresting effect of ileostomy before advanced and irreversible changes in the colon have occurred. Thus many more might fall into the first group, in whom ileostomy proves curative and the colon can be saved safely.

HARVEY B. STONE

Book Review

PRINCIPLES OF INTERNAL MEDICINE. Edited by T. R. Harrison, M.D.; Raymond D. Adams, M.A., M.D.; Paul B. Beeson, M.D., C.M.; William H. Resnik, Ph.B., M.D.; George W. Thorn, M.D., M.A.(Hon.), LL.D.(Hon.), Sc.D. (Hon.); M. M. Wintrobe, M.D., B.Sc. (Med.), Ph.D. Ed. 2. Pp. 1703; indexed. New York, 1954, The Blakiston Company, Price $16.00.

In the second edition of Harrison's *Principles of Internal Medicine*, the unorthodox approach to the subject and unconventional arrangement of the text that was inaugurated in the first edition in 1950 are continued.

The text is basically arranged in three broad sections: "The cardinal manifestations of disease," "The basic mechanisms of disease," and, finally, "A consideration of basic disease processes." Many chapters have been revised and expanded and much new material has been included. Each chapter is followed by brief, pertinent bibliographic references.

In the first section, which covers basic principles and the functional approach to the patient, the authors have made a fundamental contribution to present-day teaching. The approach to the subject is through pathologic physiology, with excellent discussions of general signs and symptoms and the differential diagnosis of underlying disease processes. Emphasis throughout is on the theme of the functional approach to the patient and his psychologic reaction to stress and disease. The authors have attempted throughout to integrate clinical medicine with the basic sciences and correlate these with psychosomatic principles of patient management. They are to be congratulated on the brilliant manner in which they have organized and presented these fundamental concepts.

A new section has been included on principles of therapy of the general problems "likely to occur in any gravely ill patient." Such varied topics as sedation and the management of abdominal distention are discussed. This discussion is very practical and is likely to prove invaluable to the intern in his early contacts with the management of hospitalized patients.

The last part of the book contains descriptive accounts of specific diseases. Many sections have been expanded and new ones added. Emphasis is on the common diseases, and the accounts of the rarer maladies are often superficial and sometimes inadequate. This is justifiable in a teaching text, where a student may familiarize himself with the common diseases without having to wade through a mire of factual minutiae, but it does seriously detract from the book's value as a source of reference.

Line drawings throughout are clear and descriptive, but the illustrations are unclear and lack sharpness of detail. The index is excellent; the paper and printing are of good quality.

This book is intended primarily for the student and young physician, and it is enthusiastically recommended by this reviewer as an introduction to internal medicine. Although any physician will find much of interest and value in the opening sections, this book lacks value to the mature physician as a reference work on the subject.

W. Beale Lutz

OL.
1
955